UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE BIOLOGIA ANIMAL

A landscape genetics approach to a contact zone of two *Salamandra salamandra* subspecies in southwest Portugal

Ana Catarina Afonso Silva

Mestrado em Biologia Evolutiva e do Desenvolvimento

2011

UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE BIOLOGIA ANIMAL

A landscape genetics approach to a contact zone of two *Salamandra salamandra* subspecies in southwest Portugal

Dissertação orientada por Doutor Carlos Alberto Rodrigues Fernandes e Prof. Doutora Maria Manuela Gomes Coelho Noronha Trancoso

Ana Catarina Afonso Silva

Mestrado em Biologia Evolutiva e do Desenvolvimento

2011

Abstract

The fire salamander (*Salamandra salamandra*) is distributed in the center and south of Europe and has a high genetic and morphological variability. In Portugal there are two subspecies described, *S. s. gallaica* with a northern and central distribution and *S. s crespoi* with a more restricted distribution in the south.

Defining the geographical boundaries between populations and subspecies is frequently difficult because there are often uncertainties on their distribution limits at putative or proven contact areas. One approach being increasingly used to detect cryptic geographic boundaries between populations is landscape genetics. This is an emerging field that integrates population genetics, landscape ecology and spatial statistics and it aims to provide information about the interaction between landscape features and microevolutionary processes.

I attempted here to determine the geographic distribution and limits of the two aforementioned subspecies at their putative contact zone in southwest Portugal, using data from microsatellites and the mitochondrial gene cytochrome-b analysed with landscape genetics tools. This approach allowed evaluating if the two subspecies are currently in allopatry and the occurrence of gene flow between them.

Results show that the two subspecies are separated by a narrow barrier that contains the Sado River and imposes a very reduced gene flow, both contemporarily and historically, between them. Currently, the most important feature of this barrier seems to be the area of sandy soil south of the Sado that was part of the riverbed of an extended river in historical times. Although both populations show signs that they have been large for some time, *S. s. crespoi* seems to have undergone a population expansion around 18,000 years BP.

Keywords: Amphibian, *Salamandra salamandra*, Subspecies, Southwest Portugal, Landscape Genetics, Mitocondrial DNA, Microsatellites, Bayesian Inferance

Resumo

A *Salamandra salamandra* ou Salamandra-de-pintas-amarelas é um urodelo pertencente à família Salamandridae, com hábitos nocturnos, sedentários e totalmente terrestres, procurando meios aquáticos apenas para se reproduzir. A espécie apresenta uma distribuição restrita no território europeu, ocorrendo nas regiões do centro e sul da Europa e rareando a norte e leste. É uma espécie altamente variável, do ponto de vista genético, morfológico, coloração e modo de reprodução. Estão descritas pelo menos 14 subespécies das quais 10 ocorrem na Península Ibérica ilustrando a importância desta região como refúgio durante os períodos glaciares. A Peninsula Ibérica actuou como um dos refúgios glaciares mais importantes da Europa. Várias evidências genéticas mostraram uma recente subdivisão deste em vários pequenos refúgios, uma vez que diferentes espécies apresentam forte subestrutura genética coincidindo com os diferentes refúgios da Peninsula Ibérica.

Steinfartz *et al.* (2000) utilizou a região controlo do ADN mitocondrial para analisar a diferenciação genética entre populações europeias. Este estudo identificou três linhagens: a linhagem A, correspondendo às subespécies *S. s. longirostris, S. s crespoi* e *S. s. morenica*, que será mais basal e terá divergido há cerca de 2 a 4 milhões de anos das restantes linhagens. A linhagem B, com populações do norte de Espanha (*S. s. bernardezi*) e do sul de Itália (*S. s. gigliolii*) que sugerem que esta linhagem terá tido uma distribuição mais ampla no passado, e a linhagem C constituída por uma politomia não resolvida das restantes subespécies.

Porteriormente García-París *et al.* (2003) utilizou o gene citocromo-*b* do ADN mitocondrial, para estudar as zonas de contacto entre subespécies na Península Ibérica com ênfase nas zonas de presença de populações ovovivíparas e vivíparas . Apesar de terem sido utilizadas poucas amostras das subespécies portuguesas, os resultados deste estudo corroboraram a existência de diferenciação entre as subespécies *S. s. crespoi* e *S. s. gallaica*.

Em Portugal estão descritas duas subespécies, *S. s. gallaica* Seone, 1884 e *S. s crespoi* Malkamus, 1983. A primeira apresenta uma distribuição ampla por todo o norte e centro do país correspondendo à linhagem mais recente identificada em Steinfartz et al. (2000), enquanto *S. s. crespoi* tem distribuição consideravelmente mais restrita, sendo um endemismo do Sudoeste de Portugal e é representativa da linhagem mais ancestral.

Os limites das áreas de distribuição das diferentes subespécies de *S. salamandra* são incertos salientando a necessidade de estudos detalhados a uma escala local para averiguar limites geográficos entre subespecies. Um estudo realizado para avaliar a distribuição de uma subespécie de Salamandra-de-pintas-amarelas no centro de Espanha detectou eventos

recentes de contacto secundário com outras subespécies. O delineamento de limites geográficos em populações e subespécies é frequentemente difícil não só devido ao problema da definição dos limites das espécies mas também por não se conhecerem as fronteiras geográficas das mesmas. A detecção de limites genéticos em espécies com distribuição contínua e onde populações discretas não são fácilmente definidas pode contribuir para a compreensão dos processos que determinaram a sua actual distribuição , como por exemplo a existência de barreiras ao fluxo genético. A genética da paisagem tem, recentemente, sido utilizada para a detecção de barreiras geográficas entre populações com distribuição continua. Esta disciplina é uma área de investigação emergente que integra fundamentos da genética populacional, ecologia da paisagem, e estatística espacial na avaliação do efeito das variáveis da paisagem e ambientais na diversidade genética e estrutura populacional.

Foram utilizados três métodos baseados em inferência bayesiana, correntemente utilizados em genética da paisagem, para identificar o limite da distribuição geográfica das duas subespécies de salamandra na zona de contacto do Sudoeste de Portugal (*S. s. gallaica e S. S. crespoi)* através do uso de 10 microssatelites já descrito para a espécies. Os resultados mostraram a separação das duas subespécies por uma barreira com cerca de 20km de extensão que integra o rio Sado e uma área envolvente predominantemente constituída por areias, arenitos e as argilas. Este resultado também foi corroborado utilizando como marcador molecular um fragmento do gene mitocondrial citocromo-*b.* Porém, para este marcador foi utilizado um menor número de amostras, que foram consideradas representativas da amostragem total utilizada para os microssatelites, minimizando a redundância de se utilizarem amostras de locais próximos. A rede de haplótipos mostrou dois grupos de haplótipos distintos para a subespécie *S. s. gallaica* na área amostrada

Apesar de alguns problemas obtidos com os loci de microssatélites utilizados (alelos nulos, desvios ao equilibrio de Hardy-Weinberg), foi possivel avaliar a ocorrência de fluxo genético entre ambas as subespécies. Com base nos resultados dos dois marcadores e de vários programas que avaliam o fluxo genético e identificam indivíduos migrantes, foi possível verificar que actualmente o fluxo entre ambas as subespecies é muito baixo.O fluxo genético reduzido no passado (detectados 3 migrantes pela análise ADN mitocondrial), parece permanecer actualmente (detectados 8 migrantes pelos microssatelites), o que sugere a existência de uma barreira forte e persistente na passagem de migrantes entre as duas populações.

Dentro de cada subespécie apenas foi detectada uma população. Com o auxilio dos dois tipos de marcadores moleculares foi possível averiguar que ambas as populações têm

apresentado ao longo do tempo um elevado efectivo. Contudo, terá ocorrido uma expansão recente em *S. s. crespoi* iniciada à cerca de 18,000 anos atrás. Esta data é coincidente com o último máximo glacial sugerindo que após este período a população de *S. s. crespoi* terá expandido a partir de um refugio no sul.

Os vários métodos utilizados para estimar o tempo de divergência das duas populações sugerem que a separação ocorreu no Pleistocénico. O método mais sofisticado, que considera qualquer sinal de fluxo genético durante a datação de uma separação populacional, sugere uma estimativa mais recente de aproximadamente à 210,000 anos. Evidências sugerem este período como quente e húmido, em que terá ocorrido expansão de florestas de *Quercus* no sudoeste da Ibéria. Para além de na Europa ter ocorrido um período semelhante por volta dos 120,000 de anos atrás, apenas no presente o clima nesta região é tão favorável para a espécie. Um período quente e húmido poderá ter originado uma expansão da rede hidrgráfica do Sado juntamente e de habitat florestal, facilitado a ocorrência de um um maior fluxo genético, que não terá voltado a repetir-se desde então.

Assim, com uma abordagem utilizada em genética da paisagem mostrou-se que no Sudoeste de Portugal as subespécies de Salamandra-de-pintas-amarelas estão separadas por uma barreira que integra o rio Sado e uma área envolvente predominantemente constituída por areias e arenitos. Para além disso, há evidências de um fluxo genético histórico e actual muito reduzido, tendo ocorrido um último período de contacto àcerca de 210,000 anos. Os resultados também evidenciam que a subespécie endémica do sudoeste de Portugal terá iniciado uma expansão àcerca de 18,000, após o último máximo glaciar.

Palavras-Chave: Anfíbios, *Salamandra salamandra*, Subespécies, Sudoeste de Portugal, Genética da Paisagem, ADN mitocondrial, Microssatélites, Inferência Bayesiana

Agradecimentos

Em primeiro lugar gostaria de agradecer ao meu orientador Carlos Fernandes que aceitou orientar-me durante este ano. Pela oportunidade de ter podido aprender a pensar e a fazer ciência, de ter podido participar em projectos diferentes do meu, por todo o apoio fornecido em qualquer momento e sobre qualquer assunto e por ter conseguido lidar com as minhas teimosias.

À Professora Maria Manuela Coelho, por também ter aceite orientar a minha tese durante este ano, mas mais que isso por ter sido a primeira a dar-me a oportunidade de desenvolver as minhas capacidades técnicas e conhecimentos científicos que foram muito importantes durante este ano. Obrigada por ter estado sempre disposta a receber-me e a ajudar-me a esclarecer as minhas dúvidas que foram surgindo ao longo destes últimos três anos.

Ao Professor Rui Rebelo, por me ter ensinado sobre este anfíbio tão interessante que é a Salamandra-de-pintas-amarelas, pelas orientações de onde procurar amostras e por ter estado sempre acessível para responder às minhas dúvidas.

À equipa do projecto landgen que todos de alguma forma contribuíram para esta tese, em especial à Jacinta por ter sido a minha irmã mais velha do laboratório (como dizia o Carlos), ter estado presente em todas as etapas desde o trabalho de campo à discussão de ideias no final, pelo apoio em todos os momentos bons e maus e pelos nossos passeios e idas às compras. Também um grande obrigado à Luciana que se não fosse por ela grande parte da amostragem não teria sido possível, e por me ter mostrado que é possível uma mulher ser uma senhora mas simultaneamente ser muito forte para fazer trabalho campo. Obrigado à Fabiana que muito ajudou na amostragem para a minha tese quando também tinha a dela para fazer. Agradeço também ao Fernando pelo contributo importante para a minha amostragem. E finalmente à Mafalda Basto por ter estado sempre lá para me apoiar, ouvir e responder às minhas dúvidas.

Agradeço também ao pessoal do grupo dos carnívoros (Mónica, Miguel, Nuno, Tátá, Diana, etc.) pelos lanches do C5 que em muitos dias stressantes era o que ajudava a desanuviar.

Ao grupo “genes, gene expression and evolution” pelas reuniões que me ajudaram a desenvolver um pouco o espírito crítico científico, em especial à Mónica, Maria Ana, Ana Rita, Tiago e Miguel M. que de alguma forma nestes últimos três anos sempre estiveram disponíveis para as minhas dúvidas.

Obrigada à Senhora Cristina Seleiro por ter se disponibilizado a imprimir esta tese. E à Senhora Vitória que foi um bom apoio nestes últimos anos.

À minha Raquelita por ter estado presente durante todo o ano e ter-se preocupado comigo, pela ajuda no campo e por ter estado sempre disponível para me ouvir. Também um obrigado pela força da Marta, Ninda e da Joana Pinho.

Às minhas meninas. À minha Inês por me mostrar o que é a ciência do ponto de vista social e humano, à minha Yola pelo apoio e pela ajuda nas imagens todas giras que apresento nesta tese, e à minha Daniela que está sempre lá para me apoiar seja em casa ou na faculdade, no telemóvel ou pela internet.

Aos meus pais que sempre trabalharam muito para me darem tudo o que conseguiam para eu chegar onde estou hoje, à minha irmã por durante este ano ter me aturado a mim e às minhas maluquices enquanto estou a trabalhar, aos meus avós e resto da família por demonstrarem interesse num bicho que até lhes faz um pouco de nojo.

Ao André, o meu melhor amigo e que me ensina tanta coisa todos os dias, obrigada por puxares sempre por mim para dar o meu melhor, obrigada por fazeres de mim uma melhor bióloga e obrigada por toda a força que foste capaz de dar quando estavas a 3000km de distância.

Contents

Preliminary note

This thesis is part of the FCT project "Effects of habitat fragmentation on the population structure and connectivity of forest-dwelling species: a comparative landscape genetics approach" (FCT PTDC/BIA-BEC/101511/2008). It aims to identify corridors that connect the two *Natura 2000* sites Cabrela and Costa Sudoeste with a comparative landscape genetics approach to four highly distinct species of forest-dwelling terrestrial animals. The study species are a carnivore (*Martes foina*), a rodent (*Apodemus sylvaticus*), an insect (*Carabus lusitanicus*) and an amphibian (*Salamandra salamandra*). Initially, the main purpose of this thesis was to carry out the part of the project on the *Salamandra salamandra*, but due to the nature of the obtained results the subject of the thesis had to be changed.

Introduction

Salamandra salamandra

The fire salamander, *Salamandra salamandra* (Linnaeus, 1758) is a urodele that belongs to the family Salamandridae, order Caudata. This species is mainly nocturnal, sedentary and completely terrestrial when adult, only search water bodies for reproduction. Live expectancy can reach up to 25 years in the wild, and reproduction typically starts after the fourth year (Rebelo & Caetano 1995). Salamanders are active after September until the end of April, having an annual activity period of roughly seven months split in two periods: the autumn rains of October/November and the late winter/early spring rains of February/March (Rebelo & Leclair 2003). Concerning environmental conditions, the species is markedly dependent of high relative humidity, absence of wind, and nocturnal temperatures between 4ºC and 14 ºC (Rebelo 2008). Salamanders find these conditions in temperate deciduous forests, although they may occur in a large variety of habitat types. The species' habitats range from subalpine meadows to Mediterranean scrubland, and even in areas of steppe when these occur close to water-courses with shrub cover. Salamanders are very effectively protected against predation since they produce a toxin in their parotid glands that repels predators, with the coloured portions of the salamander's skin being considered a warning signal to their toxicity (Rebelo 2008).

Salamandra salamandra is a species with a wide distribution and high morphological variability. Consequently, several different subspecies were described across different geographic regions. The species was historically considered to constitute a polytypic species assemblage with 16 subspecies, distributed over three continents, based on the distinctive morphology (Eiselt 1958; Klewen 1991). However, Steinfartz *et al.* (2000) were able to show, using genetic markers, that it was possible to separate the genus *Salamandra* in six different species as supported by previous studies (Joger & Steinfartz 1994; Veith 1994), corresponding to *S. salamandra*, *S. algira*, *S. infraimmaculata*, *S. corsica*, and the previously considered *S. atra* and *S. lanzai*, each with their associated subspecies. Hence, the 16 subspecies previously considered within the *S. salamandra* complex are presently classified in 13 subspecies. The study by Steinfartz *et al.* (2000) suggests the presence of three main lineages within *S. salamandra*: (A) the most ancient, comprising the subspecies *S. s longirostris*, *S. s. morenica* and *S. s. crespoi*, that comes out as a reasonably well separated unit from the remaining lineages at about 2 to 4 million years ago; (B) a second lineage, including populations from northern Spain (*S. s. bernardezi)* and populations from southern Italy (*S. s. gigliolii)*, inferred to represent the remnant

of a formerly widespread lineage that covered the whole of Central Europe in a previous interglacial period, approximately 500 000 years ago; and (C) a more derived lineage, represented by the remaining *S. salamandra* subspecies as a non-resolved polytomy, which indicates a single founder population that only relatively recently has expanded.

In the Iberian Peninsula this salamander (*S. salamandra*) is highly diversified and ten subspecies are currently recognized (Salvador, 1974; Gasser, 1978; Joger & Steinfartz, 1994; Veith, 1994; Steinfartz *et al.*, 2000): *S. s. terrestris*, *S. s. almanzoris*, *S. s. bejarae*, *S. s. bernardezi*, *S. s. fastuosa*, *S. s. gallaica*, *S. s. crespoi*, *S. s. morenica*, *S. s. longirostris* and *S. s. alfredschmidti* (described recently in Köhler & Steinfartz 2006). With the exception of the first subspecies, they are considered as endemic of the Iberian Peninsula (Salvador & García-París 2001; Köhler & Steinfartz 2006). It is currently under discussion if the subspecies *S. s. longirostris* should be regarded as a different species, considering the results of the genetics studies by Steinfartz *et al.* 2000 and García-París *et al.* 2003. These two studies are important because they both show that the genetic lineages identified in the Iberian *S. salamandra* subspecies do not corroborate the nine subspecies that have been described in this geographic area as belonging to the polytypic species as described by Eiselt (1958) and Klewen (1991). García-París *et al.* (2003) analysed sequence variation in the mitochondrial gene cytochrome-*b* (cyt*b*) in populations sampled throughout the Iberian Peninsula, and investigated the patterns of haplotype diversity within the three contact zones between the two viviparous subspecies (*S. s. bernardezi* and *S. s. fastuosa*, which progeny consists of relatively developed small larvae that are directly released into streams and ponds) and the ovoviviparous surrounding populations (*S. s. gallaica* and *S. s. terrestris*). The subspecies *S. s. longirostris* was found not to belong to the same clade that the rest of the Iberian subspecies and the latter were divided into five distinct genetic lineages. Like in Steinfartz *et al.* 2000, *S. s. crespoi* and *S. s. morenica* are considered within the same lineage, whereas *S. s. almanzoris* appears as a separated lineage but still within the same clade. A third lineage corresponds to the polytomy of the subspecies *S. s. gallaica*, *S. s. terrestris*, *S. s. bejarae* and the ovoviviparous populations of *S. s. fastuosa*. The other two lineages represent the viviparous populations of *S. s. fastuosa* and *S. s. bernardezi*. García-París *et al.* (2003) also showed discordances between classifications based on external morphology (this is the criteria currently used to recognize subspecies), nuclear DNA (allozyme studies; Alcobendas *et al.* 1996) and mitochondrial DNA (mtDNA) data, along two of the three contact zones in northern Spain. These discordances are a consequence of hybridization after secondary contact of previously isolated lineages (García-París *et al.* 2003).

The current distribution of mtDNA genetic variation among widespread temperate populations of different taxa is subdivided into patches of vicariant allospecies, subspecies, or ecotypes. The climatic oscillations of the Pleistocene led to range reductions, population extinctions, and isolation in southern refugia during glacial periods, followed by postglacial range expansions that generated parapatric distributions and secondary contact zones currently observed (e.g. Bernatchez & Wilson 1998; Hewitt 1999; Zamudio & Savage 2003; Gómez & Lunt 2007). The Iberian Peninsula was one of the most important Pleistocene glacial refugia in the European continent (Hewitt 1999; Hewitt 2001). There is growing genetic evidence for recent subdivision among multiple refugia since many species display a strong population substructure within the Iberian glacial refugium, with the reciprocal monophyly of phylogeographic lineages suggesting physically-separated histories throughout the Pleistocene. It is this lack of admixture between phylogeographic lineages, portrayed by geographically structured clades, which points to distinct Pleistocene 'refugia-within-refugia' (Gómez & Lunt 2007). Such instances are seen in animals and plants such as *Squalius* (Brito *et al.* 1997; Coelho *et al.* 1998; Zardoya & Doadrio 1998), *Discoglossus* (García-París & Jockusch 1999), *Chioglossa lusitanica* (Alexandrino *et al.* 2000), *Lacerta schreiberi* (Paulo *et al.* 2001; Paulo *et al.* 2002), *Natrix maura* (Guicking *et al.* 2008), *Oryctolagus cuniculus* (Branco *et al.* 2000; Branco *et al.* 2002), *Arvicola sapidus* (Centeno-Cuadros *et al.* 2009), *Elona quimperiana* (Vialatte *et al.* 2008), *Artemia salina* (Muñoz *et al.* 2008), and *Quercus* (Olalde 2002).

In Portugal, the fire salamander has a broad distribution but in the south it occurs within relatively small habitat patches, being absent in agricultural areas in the Baixo Alentejo region due to the lack of tree cover (Rebelo 2008). Climatic and habitat features shape the distribution of the two subspecies described in Portugal: *S. s. gallaica* (Seone, 1884) and *S. s crespoi* (Malkamus, 1983). The first corresponds to the most recent lineage in Steinfartz *et al.* (2000) and has a wider distribution, from the north of Portugal to the Alto Alentejo region. It is characterized by smaller adults with a predominantly dark colour, few large rounded yellow spots, and the occasional presence of red spots on the flanks and throat. *S. s. crespoi* belongs to the most ancient lineage in Steinfartz *et al.* (2000) and presents a more restricted distribution, being a Southern endemism in Portugal. This subspecies is characterized by a dorsal pattern with numerous small and irregular yellow dots that often show an elongated form (Rebelo 2008; Almeida *et al.* 2001). Although specimens from Portugal were included in the studies by Steinfartz *et al. 2*000 and García-París *et al. 2*003, only 2 and 3 individuals were analysed, respectively, from the areas in which the two subspecies may contact. Besides, in one study (Steinfartz *et al.* 2000) a single sample from Alentejo was assigned to *S. s. gallaica*, whereas in

the other study (Garcia-Paris *et al.* 2003) a sample from the same region was assigned to *S. s. crespoi*.

In fact, the limits of the distribution areas of the different *S. salamandra* subspecies remain mostly uncertain, emphasizing the need of more detailed studies at the spatial scale of the postulated geographic boundaries between subspecies. Martínez-Solano *et al.* (2005) carried out a genetic survey, using the gene cytochrome-*b* and allozymes, on the endangered subspecies *S. s. almanzoris*, which is the subspecies with a more restricted distribution in the Iberia Peninsula. They were able to delimit the distribution area of the subspecies, but found discordance between the results of the different genetic markers, the coexistence of divergent mtDNA haplotypes in two populations, and discordance between external morphology and mtDNA haplotypes in different populations. These were suggested to be consequence of recent events of secondary contact between *S. s. almanzoris* and other subspecies of *S. salamandra.*

Landscape genetics

Defining the geographical boundaries between populations and subspecies is frequently difficult because, as with the definition of species limits (but even more so), there are often areas of indeterminacy at prospective borders (Rowe & Beebee 2007). The detection of genetic boundaries in systems where species are either continuously or patchily distributed, and where discrete populations are not easily defined, may help to determine the underlying generating processes, such as important historical events (Miller *et al.* 2006) or current barriers to gene flow (Coulon *et al.* 2006; Cushman *et al.* 2006; Safner *et al.* 2010). One approach being increasingly used to detect geographic boundaries between populations is landscape genetics (Manel *et al.* 2003).

The discipline of landscape genetics is an emerging field that integrates population genetics, landscape ecology and spatial statistics (Manel *et al.* 2003; Storfer *et al.* 2007; Holderegger & Wagner 2008) and its original definition states that it aims to provide information about the interaction between landscape features and microevolutionary processes (gene flow, genetic drift and selection) (Manel *et al.* 2003). It uses spatially explicit models to examine how landscape features affect the spatial distribution of genetic variation (Manel *et al.* 2003; Holderegger & Wagner 2006; Storfer *et al.* 2007).

Storfer *et al.* (2010) reviews the several questions that have been addressed in landscape genetics: identification of genetic discontinuities, cryptic population boundaries, and

specific barriers to dispersal (Manni *et al.* 2004; Funk *et al.* 2005; Zannèse *et al.* 2006; Rowe & Beebee 2007; Latch *et al.* 2008; Gilles Guillot 2008), quantifying genetic diversity (Domínguez-Domínguez *et al.* 2007; Segelbacher *et al.* 2008), inferring the effect of landscape change (Wagner *et al.* 2006; Wang *et al.* 2008), identification of migrants in relation to landscape condition (Wilmer *et al.* 2008; Janssens *et al.* 2008), estimation of source-sink dynamics (Lowe *et al.* 2006; Martínez-Solano & González 2008), prediction of spread of disease (Foley *et al.* 2005; Wood *et al.* 2007; Deter *et al.* 2008) or invasive species (Lecis *et al.* 2007), and comparing observed genetic patterns with contemporary and historic landscapes (Orsini *et al.* 2008; Spear & Storfer 2008). Landscape genetics is also useful to study gene flow between populations, subspecies and species (Fitzpatrick & Shaffer 2007; Fitzpatrick *et al.* 2008; Dionne *et al.* 2008; Maletzky *et al.* 2010; Leaché 2011; Bolfíková & Hulva 2011), for instance to assess the presence of admixture or hybridization in zones of secondary contact.

The landscape genetics approach does not require an a priori identification of discrete populations and considers the individual as the operational unit of analysis within the population (Manel *et al.* 2003; Wright 1931). Landscape genetics usually focuses in relatively small spatial scales, where individuals of a study species are mostly closely related; hence its laboratory analysis has to rely on highly variable molecular markers that provide enough resolution power (Holderegger & Wagner 2008). Currently, two different types of genetic markers are dominant in landscape genetics, namely amplified fragment length polymorphisms (AFLPs) and single sequence repeats (SSRs), with the latter being the mostly used (Storfer *et al.* 2010). Landscape genetics uses several different methodological tools to identify genetic patterns: Wright's F_{ST}, Mantel's test, spatial autocorrelation, ordination analyses and synthesis maps, Bayesian clustering approaches, edge detection methods (Monmonier's algorithm and Wombling) and other individual-based methods (e.g. Manel *et al.* 2007; Murphy *et al.* 2008).

Two general families of methodologies are currently used to identify boundaries in landscape genetics: edge detection methods and Bayesian clustering algorithms (Safner *et al.* 2011). Edge detection methods identify areas where changes in variables occur from the analysis of allele frequency data, using approaches such as the Monmonier's algorithm (Monmonier 1973) and wombling (Womble 1951; Barbujani *et al.* 1989). Bayesian clustering algorithms intend to identify discrete sets of individuals based on the analysis of multilocus genotypes (Pritchard *et al.* 2000; Corander *et al.* 2003). The softwares that use these algorithms delineate clusters of individuals based on the analysis of individual genotypes, and also assign individuals to the identified cluster where their posterior probability is highest. In spatial models the probability that two individuals belong to a same cluster is influenced by their geo-

graphic distance, whereas geographic proximity is ignored in non-spatial models (Safner *et al. 2*011). In a recent comparison, Bayesian spatial clustering algorithms outperformed edge detection methods, although all methods incorrectly detected boundaries in the presence of strong patterns of isolation by distance (Safner *et al. 2*011).

Aims

With this work I intended to a) determine the geographic distribution and limits of the two Portuguese *Salamandra salamandra* subspecies at the putative contact zone in the southwest of the country, using data from the mitochondrial gene cytochrome-*b* and microsatellites analysed with landscape genetics tools; b) investigate if the two subspecies are currently in allopatry as currently accepted in Portugal, or if there is evidence of simpatry between the subspecies; c) evaluate the occurrence of gene flow between the named subspecies; d) characterize the population genetic diversity of *S. salamandra* in the study area.

Material and methods

Tissue samples were collected from 102 adults, using either toe clippings from live animals or muscle material from road kills, and from the tail tips of 116 larvae collected at 69 different spawning sites, in a total of 216 samples for which the exact GPS coordinates were recorded. Tissue was stored in DMSO-salt solution (Seutin *et al.* 1991). The DNA was extracted using the commercial kit EZNA Tissue DNA kit (Omega Bio-tek, Inc) following the manufacturer's instructions.

Microsatellite genotyping and cytochrome-*b* sequencing

DNA extracts were screened for 10 microsatellite loci: Sal E2, Sal E6, Sal E7, Sal E8, Sal E11, Sal E14, Sal 3, SST-A6II, SST-B11 and SST-C2 (Steinfartz *et al.* 2004, Hendrix *et al.* 2010). For all loci, the forward primers were labelled with a 5' fluorescent tail (6-FAM and HEX) for visualization. Each 10-µL PCR reaction contained 50-100 ng of genomic DNA, 1× reaction buffer, 2 mM $MgCl_2$, 0.2 mM of each dNTP, 0.5 µM of each primer, 1 µg BSA (Bovine Serum Albumin), and 0.75 U of the "hot-start" HotSurf Taq DNA polymerase (StabVida). Thermal cycling consisted of a "step down" protocol, with a first part with annealing temperature at 65ºC-60ºC (depending of the primer) for 15 cycles followed by a second part with 60ºC-55ºC of annealing temperature for 30 cycles. The loci were amplified individually and the PCR products ran in sets of two loci on an ABI 310 Capillary Electrophoresis Genetic Analyzer (Applied Biosystems). Fragments were sized with ROX-500 size standard, collected with GeneScan version 3.1, scored with Genemapper version 3.7 (Applied Biosystems), and the binning was performed with TANDEM version 1.08 (Matschiner *et al.* 2009).

In Steinfartz *et al.* 2000 they used a D-loop 738 bp fragment, where it is shown that in Salamander species D-loop is not as informative as in mammals since has a low divergence and also that the whole mitochondrium DNA seems to have the same substitution rate. While in García-París *et al.* 2003 is used a small fragment (372 bp) of cytochrome-*b* gene, hence in this work it was preferred to use also cytochrome-*b* gene but new primers were designed to amplify a longer fragment that would provide more information since it is used fresh samples.

Amplification and sequencing of 702 base pairs (bp) of the mitochondrial gene cytochrome-*b* were performed for a subset of 64 samples. The primers forward SALF 5'-CCTGAAGTAGGAACCAGATG-3' and reverse SALR 5'-TCAAAGCTTACACCAGTCTTGTAAACC-3' were designed by aligning complete cytochrome-*b* sequences available in GenBank for several

species of amphibians. A 20-µL PCR reaction contained 50–200 ng of DNA, 1× reaction buffer, 2 mM $MgCl_2$, 0.2 mM of each dNTP, 0.2 µM of each primer, 2.2 µg BSA, and 1.25 U of SurfTaq DNA polymerase (StabVida). PCRs consisted of 35 cycles with a denaturing temperature of 94 °C (30 sec), annealing temperature at 57 °C (45 sec) and extension at 72 °C (1 min). Double-stranded amplification products were purified with ExoSAP (GE Healthcare Life Sciences). Purified products were sequenced in both directions using the PCR primers and the Ez-Seq DNA sequencing service at Macrogen Inc., Europe. Cytochrome-*b* sequences were aligned using SEQUENCHER version 4.7 (Gene Codes Corporation).

Microsatellites and mitochondrial DNA data analysis

The software CREATE (Coombs *et al.* 2008) and PGDSPIDER (version 2.0.0.3, available from http://www.cmpg.unibe.ch/software/PGDSpider/) were used to facilitate input file preparation for all the software used for microsatellite and mtDNA data analysis.

The program ML-RELATE (Kalinowski *et al.* 2006) was used to calculate maximum likelihood estimates of relatedness and the likelihood of four pedigree relationships between each pair of individual samples: uncorrelated, half-sibs, full-sibs and parent–offspring. ML-RELATE uses maximum likelihood estimates of the frequency of null alleles in all calculations when null alleles are present. It calculates the likelihood of each relationship for each pair of individuals and generates a matrix of relationships with the highest likelihood, as well as a confidence set for the relationship between pairs of individuals such that relationships excluded from the confidence set can be ruled out as unlikely.

To infer the number of populations in the sample set and to assign individual samples to these groups, three Bayesian clustering techniques were employed. First, I used the GENELAND software (version 3.3.0; Guillot *et al.* 2005b), which has been developed to optimize the delineation of subpopulations by incorporating spatial coordinates for each sample into the model. I varied K (the prior value for the number of subpopulations) from 1 to 10, using the matrix of genotypes, spatial coordinates for each individual (uncertainty was set at 0 as I gave a negligibly different coordinate to larvae samples taken from the same pond), and 10,000 stored MCMC iterations (1,000,000 iterations, thinning of 100) that were run 5 times. Allele frequencies were drawn from the uncorrelated frequency model that uses independent Dirichlet distributions (Pritchard *et al.* 2000), as this model has been shown to perform better than the alternative model (F-model; (Guillot *et al.* 2005a). The maximum rate of the Poisson process, which controls the number of polygons in the geographical area under study, was set

to 100. This value corresponds to strongly fragmented partitions and weak dependence on the spatial organization of populations. In the spatially-explicit GENELAND model, subpopulations are assumed to be partitioned, by Voronoi tessellation, into a union of a random number of polygons that are randomly assigned to one of *K* possible spatial clusters (Dupanloup *et al.* 2002). In this software the model is of the type free Voronoi tessellation, as the polygons are constructed independently of the sampling sites and are developed with basis on a continuous Poisson point process. Hence, the maximum number of nuclei within this tessellation was set to 300 (3 x maximum rate as suggested by Guillot *et al. 2*005a). For all the runs, the posterior probability of subpopulation membership was computed for each pixel of the spatial domain (100 x 100 pixels), using a burn-in of 200 iterations.

Second, I used the software BAPS 5.4 (Corander & Marttinen 2006; Corander *et al.* 2008), which treats the number of populations as an unknown parameter that is directly inferred from the data set without defining a prior estimate and also uses Voronoi tessellation but in this case it is based on discrete sampling sites. For inferring population structure in the mitochondrial (cytochrome-*b*) and nuclear data sets, the individuals were assigned using the models "clustering of linked loci" (Corander & Tang 2007) and "spatial clustering of individuals", respectively. The first does not take into account spatial information while the second does. The software was run with a predefined maximum of *k* ranging from 1 to 6 and runs were repeated five times in order to check the stability of the results. The same clusters were assigned in all independent runs. The admixture analyses were performed based on the results of microsatellites mixture clustering, using 500 iterations and a number of 1000 reference individuals per population, each with 10 reiterations.

Finally, I used the software STRUCTURE (version 2.3.3; Pritchard *et al.* 2000) to infer genetic structuring by assigning individuals to a number of genetic clusters (*k*) without prior knowledge of their affinities. The software defines clusters within which the loci are at Hardy-Weinberg equilibrium (HWE) and linkage equilibrium. This approach considers all clustering solutions to be equally likely and does not consider the geographic coordinates. This information can however be used, after the data is processed, as a method to verify subpopulation membership in the output generated by the software. The best *K* is chosen by comparing the posterior probabilities for different K values using methods that are not free of some subjectivity. A typically used method is the one described by *Evanno et al.* (2005) but, according to Waples & Gaggiotti (2006), this method brings little improvement over previous ones. Using the options to ignore population affiliation when clustering individuals, assuming independence among loci and allowing admixture (thereby estimating the proportion of individuals with

ancestry in each cluster), I ran five independent runs of 1,000,000 iterations (following a burn-in period of 100,000) for each value of *k* from 1 to 6 (Pritchard *et al.* 2000). The online software program STRUCTURE HARVESTER (version 0.6.7; Earl & VonHoldt 2011) was used to calculate delta *k*.

The subsequent analyses were done considering the clusters that were defined by these Bayesian clustering software programs.

In order to detect, within each population, genotyping errors due to the presence of null alleles, stuttering or large allele dropout, I used the bootstrap approach implemented in MICROCHECKER 2.2.3 (Van Oosterhout *et al.* 2004). Additionally, a maximum-likelihood estimate of the frequency of null alleles was calculated using the software FREENA (Chapuis & Estoup 2007), which uses the algorithm of Dempster *et al.* (1977) and has been shown (Chapuis & Estoup 2007) to perform better than the other null allele frequency estimators available in MICROCHECKER. Estimates of F_{ST} and Cavalli-Sforza and Edwards' (1967) genetic distance were also calculated using, respectively, the correction methods ENA and INA proposed by Chapuis & Estoup (2007), and that have been shown to efficiently correct the positive bias induced by the presence of null alleles on these measures and provide accurate calculations in presence of null alleles.

Three-dimensional factorial correspondence analysis (FCA) was performed to visualize patterns of differentiation between populations using the GENETIX program (version 4.05.2; Belkhir *et al.* 2004). The result of the FCA represents the relationship between individuals based on the detection of the best linear combination of allele frequencies.

To assess migration, three approaches were used: the BAYESASS software (version 1.3; Wilson & Rannala 2003) that estimates current migration rates among populations, the USEPOPINFO model in STRUCTURE that uses the sampling locations to test for migrants, and the GENECLASS software (version 2.0; Piry *et al.* 2004) that has the 'detection of first generation migrants' function explicitly designed to identify first generation migrants (Paetkau *et al.* 2004). BAYESASS relies on multilocus genotypes and a Markov Chain Monte Carlo (MCMC) algorithm to estimate proportions of non-migrants as well as the source of migrants for each sampling site (Wilson & Rannala 2003). This software is a non-equilibrium Bayesian method that does not require data sets to conform to HWE proportions. I performed five independent replicate runs of the algorithm for 9,000,000 iterations with 3,000,000 iterations discarded as burn-in. Delta values of migration rate (Δm), allele frequencies (Δp) and inbreeding coefficient (ΔF) were, respectively, $\Delta m = 0.15$, $\Delta p = 0.15$, and $\Delta F = 0.15$, which yielded an average number

of changes in the accepted range. I performed the additional round of admixture clustering with STRUCTURE, now indicating the population of origin for each individual and setting the prior probability of each individual having pure ancestry from its assigned population at 0.95 (USEPOPINFO option, MIGRPRIOR = 0.05), and the other parameters as described above. This model assumes that most individuals have pure ancestry but that a proportion of individuals may also have ancestry from other populations. GENECLASS uses several likelihood-based statistics, in combination with resampling methods, to calculate the first generation migrant probabilities. I used the L= L_home/L_max likelihood computation wherein L_home is the ratio of the likelihood of the individual genotype within the population where the individual has been sampled and L_max the highest likelihood value among all available population samples (Paetkau *et al.* 2004). To detect the first generation migrants, I used the Bayesian criterion of Rannala & Mountain (1997), in combination with the resampling method of Paetkau *et al.* (2004) to determine the critical value of the test statistic beyond which individuals were assumed to be migrants. I selected an alpha level of 0.05 to determine critical values and 10,000 as the number of simulated individuals.

MSA (version 4.05; Dieringer & Schlötterer 2003) and GENEPOP (version 4.1, Rousset 2008) were used to estimate genetic diversity in each obtained cluster as measured by the parameters number of alleles per locus (N_A), allelic richness (A_R), observed (H_o) and unbiased expected (H_e) heterozygosities, and to test for Hardy–Weinberg equilibrium (HWE) and for linkage disequilibrium (LD) between loci. The HWE was verified testing for a deficiency of heterozygotes relative to Hardy–Weinberg expectations for each locus, using the Markov chain method for estimating P-values (Guo & Thompson 1992; parameter values of dememorization = 50,000,000, number of batches = 100, number of iterations per batch = 1,000,000). The same method was used to assess LD (parameter values of dememorization = 100,000,000, number of batches = 200, number of iterations per batch = 50,000,000). A sequential Bonferroni correction (Rice 1989) was used to control for multiple comparisons.

Spatial structure was examined using a Mantel test, performed with GENALEX (version 6.0; Peakall & Smouse 2006). The Mantel test was performed between the triangular matrix of pairwise geographical distances between individuals and the triangular matrix of pairwise ϕ_{st} values between individuals.

I used the software BOTTLENECK (version 1.2.02; Piry *et al.* 1999) to evaluate if there were signs of a genetic bottleneck signature in the microsatellite data. In particular, I carried out 100,000 replications using the infinite alleles model (IAM) and the stepwise mutation

model (SMM). The sign test compares the number of loci that present a heterozygosity excess relatively to the number of such loci expected by chance alone.

For the analyses of mitochondrial DNA, I used a subset of 64 samples from the total sample set (N=203) that attempted to maintain the level of geographic coverage of the total sample set while minimising the eventual redundancy of using multiple samples from the same sampling sites (the used samples are identified in Appendix table 1).

The median-joining algorithm in the software NETWORK (version 4.6.0.0; Bandelt *et al.* 1999) was used to infer a network of haplotype relationships. Information on the geographic distribution of the haplotypes was subsequently added to the network using a colour code to represent geographic locations.

Genetic diversity for the cytochrome-*b* fragment was measured in each population by the number of haplotypes, haplotype diversity (*h*) and nucleotide diversity (π) as estimated by DNASP (version 5.10.01; Librado and Rozas 2009). ARLEQUIN (version 3.5.1.2; Excoffier *et al. 2*010) as used to estimate mean population pairwise F_{ST} statistics that were computed using the model of Tamura and Nei (1993). Significant departures from the null hypothesis (no genetic differentiation) were tested using 10,000 permutations.

The software ARLEQUIN was also employed to carry out two neutrality tests known for their additional ability to detect signatures of historical population expansion. The tests are Tajima's D (Tajima 1989a, 1989b, 1993), which compares the number of segregating sites in the sample and a parameter based on the mean number of pairwise differences between haplotypes, and Fu's F_S (Fu 1997), based on the probability of observing *k* alleles in a sample of a given size conditioned on the observed average number of pairwise differences. Also in order to test for demographic expansions, statistics based on the mismatch distribution were estimated with ARLEQUIN. The mismatch distribution (MMD) is the observed distribution of pairwise differences between pairs of haplotypes. This distribution is usually multimodal in samples drawn from populations at demographic equilibrium, but it is usually unimodal in populations having passed through a recent demographic expansion (Rogers & Harpending 1992; Slatkin & Hudson 1991), or through a range expansion with high levels of migration between neighbouring demes (Ray *et al.* 2003; Excoffier 2004). The observed distribution of pairwise differences was calculated within each population and compared with the expected results under a sudden-demographic and a spatial-demographic expansion model. Statistically significant differences between observed and simulated expected distributions were evaluated using the sum of the square deviations (SSD) and Harpending's raggedness index (RAG) (Harpending

1994). The time since the expansion began was estimated using the MMD τ (tau, estimator of the time since expansion) with $\tau = 2\ n\ \mu\ TGt$, where n is the number of nucleotides of the DNA sequence, μ is the substitution rate in mutations per site per year and TGt is the time in generations (Rogers & Harpending 1992).

Population divergence time was estimated using the coalescent method described by Gaggiotti & Excoffier (2000), also implemented in ARLEQUIN. The method removes the effect of bottlenecks and unequal sizes of the derived populations, which can lead to the overestimation of divergence times from genetic distances. Here τ (tau) is also the time estimator from genetic distance but its conversion to years before present is done with a slightly different formula: $\tau = 2\ n\ \mu\ T$, where n is the number of nucleotides of the DNA sequence, μ is the substitution rate in mutations per site per year and T is time in years. Divergence time was also estimated using a similar equation $D_A = 2\ n\ \mu\ T$ in which n is the number of nucleotides of the DNA sequence, μ is the substitution rate in mutations per site per year, T is the divergence time, and D_A is the net number of nucleotide differences between populations (Nei & Li 1979).

Population divergence time was also estimated using the software IMA (Hey & Nielsen 2007), which implements a coalescent-based isolation with migration model and can be applied to genetic data drawn from a pair of closely related populations or species (Nielsen & Wakeley 2001) to infer some demographic parameters: effective population sizes of the present and of the ancestor populations, migration rates between the daughter populations, and the splitting time (t). The program performs a Markov chain Monte Carlo (MCMC) simulation with the Metropolis-Hastings algorithm, in which random samples from the joint posterior probability density of the model parameters are generated. To evaluate if the program has run for enough time, so that the sample has converged to the true posterior probability distribution, the effective Sample Size (ESS), which is a measure of the independence of the recorded values over the course of the run, should have values greater than 200 for any estimated parameter. Four replicate simulations were run after the preliminary runs to optimize settings. Simulations used 10 Markov chains, with 45 chain swap attempts per step, were run for 50 million steps with the first 100,000 steps discarded as "burn-in", and the genealogies were sampled every 100 steps. To convert coalescent times to years before present, I used a generation time of 5 years (Rebelo 2002) and a 0.8% per MYR substitution rate. Steinfartz *et al.* (2000) used this rate value, and stated that in salamanders the substitution rate seems overall uniform across the mitochondrial DNA, and studies in other urodeles have used similar rates (Tan & Wake 1995; Caccone *et al.* 1997; Mueller 2006).

Results

Microsatellites

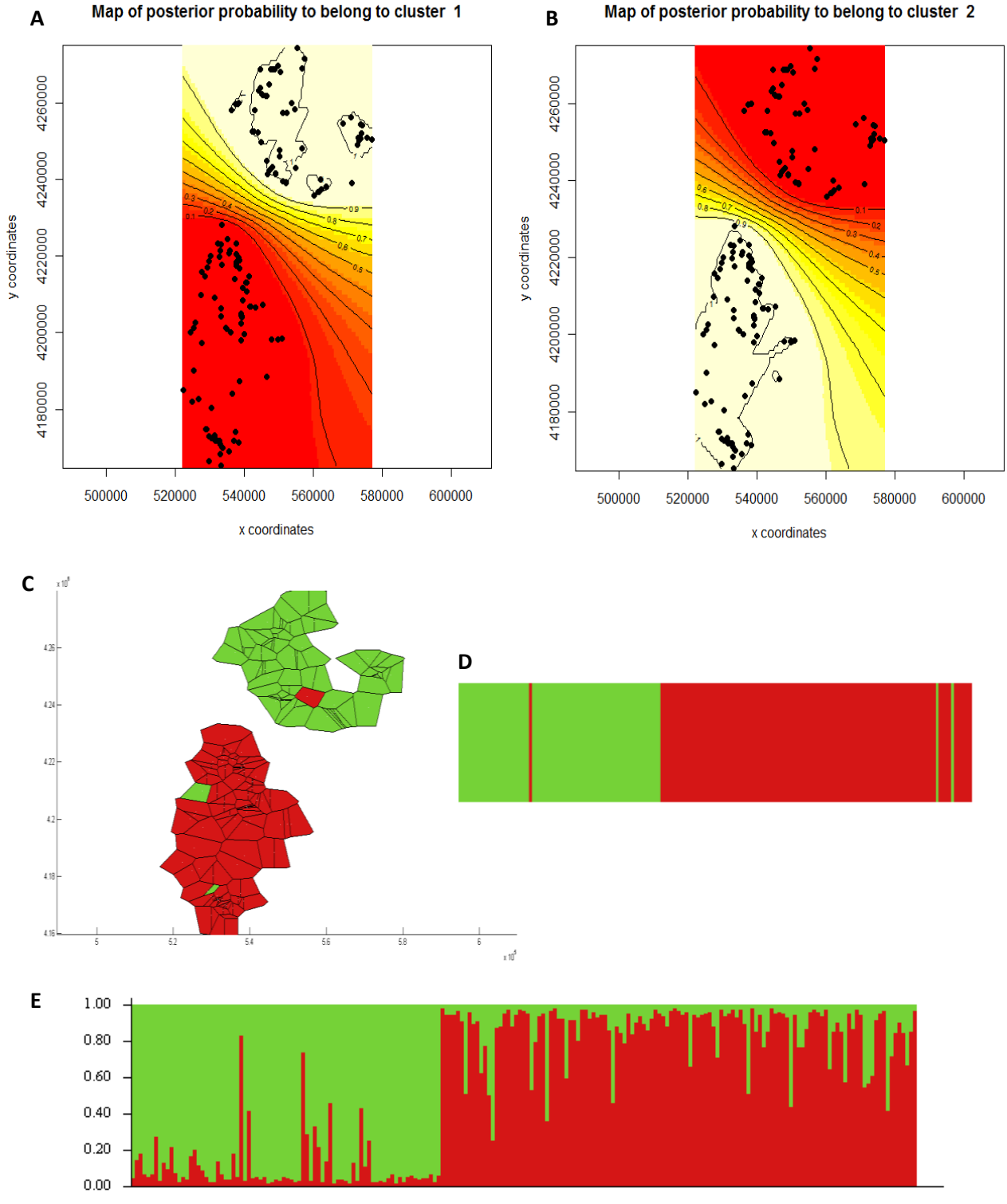


Fig. 1 Bayesian clustering results based on 10 microsatellite loci. A and B: maps of posterior probability for each cluster accordingly with GENELAND. These maps are a tessellations analysis representation with the probability that a sample belongs to a particular group ranging from low (dark colour) to high (light colour). C and D: results from BAPS; C is a spatial Voronoi tessellation representation of the admixture plot (D) in which each bar represents an individual. Northern population represented by green and Southern by red. E: STRUCTURE plot for the best-supported *k* (=2) as inferred with the method of Evanno *et al.* (2005); each bar represents an individual. Northern population represented by green and Southern by red.

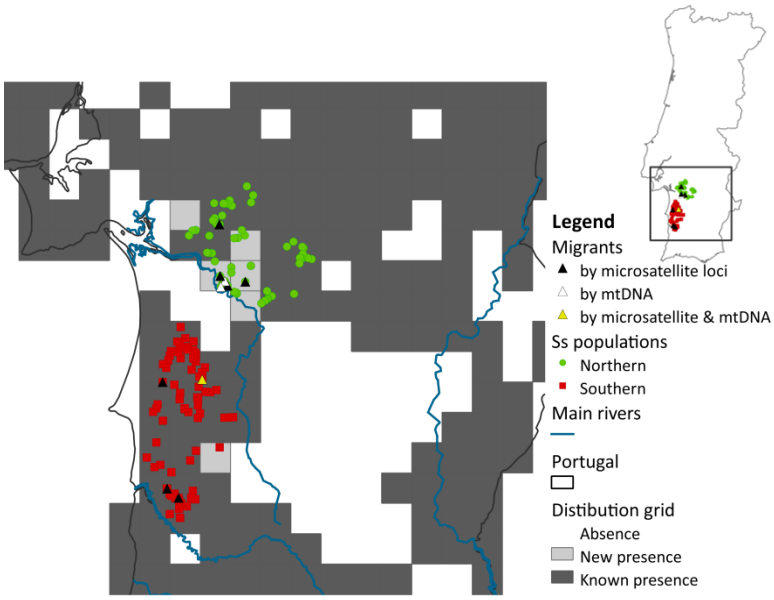


Fig. 2 Distribution map of *Salamandra salamandra* according to Rebelo (2008) and sample locations within the postulated area of two subspecies contact in southwest Portugal.

Samples were collected between 2008 and 2010 from road kills, live adults, and larvae from ponds and small streams. During the fieldwork it was detected the presence of *Salamandra salamandra* in 6 new 10 x 10 km UTM squares comparing the actual known distribution (Rebelo 2008). The 216 samples were evaluated with the software ML-RELATE to remove all full-sibling and parent-offspring samples that could be present in the sampling. Goldberg & Waits (2010) have shown that the removal of these relatives reduces the bias of sampling larvae and improves estimates of population genetic parameters in studies of amphibians. The full-siblings in a distance of approximately 1 km were removed. Three adults and 10 larvae were excluded, so the data set for subsequent analyses concerned 203 samples from 168 different sampling locations (fig. 2). All analyses were done with the samples ordered by geographical Y axes of their coordinates and no samples with missing data were taken into account.

All the 10 loci were described as tetranucleotide microsatellite loci (Steinfartz *et al.* 2004 and Hendrix *et al.* 2010), but in the populations analysed here the loci Sal E2 and Sal E11 behaved like dinucleotide loci in terms of allele bins.

All the three software programs used to infer the number of clusters indicated the presence of 2 clusters (fig. 1). In GENELAND the samples were consistently assigned to the same cluster in all runs, even if the density for the best *k* was not much higher than 40% (Appendix fig. 1). In figures 1A and 1B it is visible that all samples were assigned with high probabilities to each cluster and how the two clusters are spatially distributed. BAPS assigned almost all samples to the same two clusters that GENELAND detected. Because the spatial clustering of individuals (Corander *et al.* 2008) was used, the output of BAPS gives a spatial representation of the clustering. This allowed to identify three samples that should be migrants or admixed individuals (figures 1C and 1D).

STRUCTURE also identified *k* = 2 (fig. 1E) as best fitting the data, using the method of Evanno *et al.* (2005), but with more admixture that is probably due to the software sensitivity when the assumptions are not met (see below). Taking into account the findings of Safner *et al.* (2011), the clustering solution offered by GENELAND was considered the most reliable. The geographic boundary between the clusters seems to match with an area south of the River Sado in which no fire salamanders could be found despite being prospected several times in different times, seasons and years, for both adults and larvae. The distribution map with the samples assigned to each of the two clusters is presented in fig. 2. Given that the geographic location of the clusters correspond to the areas north and south of the detected boundary they are henceforth designated, respectively, as "Northern" (80 individuals) and "Southern" (123 individuals).

After the population genetic structure of the data was identified, the presence of null alleles in the analysed microsatellite loci was tested using two methodologies: the algorithm of Oosterhout *et al.* (2004), incorporated in the software MICROCHECKER, and the algorithm of Dempster *et al.* (1977) integrated in the software FREENA. The second estimator was chosen because it has been shown to perform better than the other estimators available in MICROCHECKER (Chapuis and Estoup 2007). MICROCHECKER found no evidence for scoring errors due to stuttering or large allele dropout, but found evidence for null alleles in some of the loci.

Table 1 Estimation of null allele frequencies by population using the algorithm of Oosterhout *et al.* (2006), as implemented in MICROCHECKER, and the algorithm of Dempster *et al.* (1977) as implemented in FREENA.

	Northern		Southern	
Locus	Oosterhout	Dempster	Oosterhout	Dempster
Sal 3	-0.0147	0.00002	0.0506*	0.04703
Sal E2	0.1202*	0.12062*	0.0584*	0.06137*
Sal E8	0.0325	0.04598	0.0328	0.03109
Sal E11	0.0405	0.04093	0.0347	0.03288
Sal E14	0.0577*	0.06378*	0.1954*	0.17553*
Sal E6	-0.0292	0.00000	0.0132	0.00696
Sal E7	0.009	0.00000	0.0562*	0.04051
SST-A6II	0.0348	0.04143	0.0028	0.00000
SST-C2	0.0143	0.02016	0.0354	0.01719
SST-B11	0.2102*	0.18448*	0.0162	0.02630

* Indicate null allele loci and grey rows indicate loci with null alleles across populations

Among the 10 analysed loci, only two consistently exhibited signs of the presence of null alleles with both approaches and in both clusters (table 1). Given the results of MICROCHECKER, the Bayesian clustering analyses to infer the number of clusters were repeated using only the eight loci not showing signs of the presence of null alleles, and the results did not visibly differ from the analyses using the 10 loci (Appendix fig. 1). Chapuis and Estoup (2007) described a method that allows estimating the F_{ST} following Weir (1996) and the Cavalli-Sforza and Edwards' (1967) genetic distance from microsatellite data sets harbouring null alleles. The corrected F_{ST} between the two clusters was 0.034309 and the uncorrected one was 0.034998, while the cor-

rected Cavalli-Sforza and Edwards' genetic distance was 0.338510 and the uncorrected one was 0.329332. This shows that, despite the presence of null alleles in two of the 10 analysed loci, analyses using the whole data set do not seem to be much affected by the presence of null alleles, and thus the subsequent analyses were always done using all of the 10 loci.

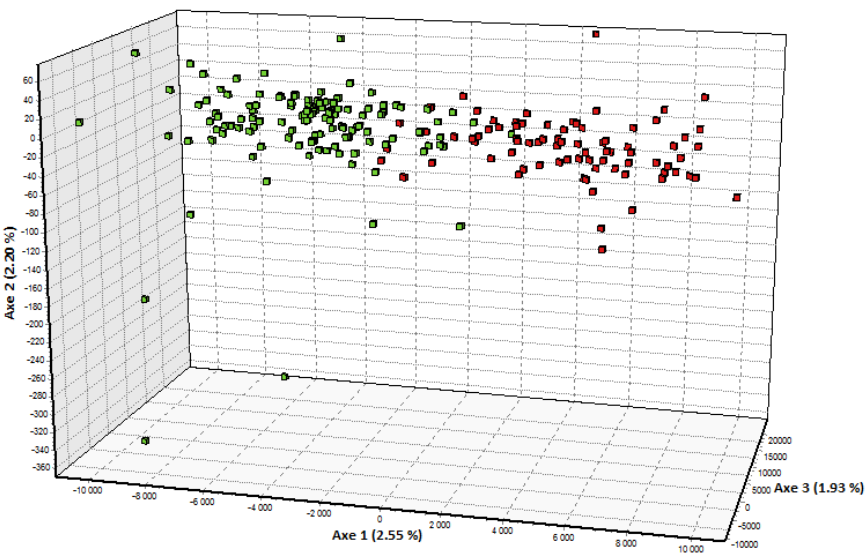


Fig. 3 Three-dimensional factorial correspondence analyses (FCA) of the salamander samples considering the two populations inferred by the Bayesian methods. Northern population represented by green and Southern by red

Factorial correspondence analysis also identified two clusters (fig. 3) in a three-dimensional space without a priori information on population structure. The axes 1 to 3 explain 6.68% of the variation among the two clusters and the two clusters appear close to each other. The result also indicates that some individuals might be genetically admixed or migrants.

Table 2 Migration rate (*m*) between the two clusters, along with their 95% confidence intervals in parentheses, estimated using BAYESASS (Wilson and Rannala 2003) from initial conditions of Δm , Δp and ΔF = 0.15.

Source	Migration rate (*m*) to	
	Northern	Southern
Northern	**0.9924±0.0072 (0.9723, 0.9998)**	0.0052±0.0048 (0.0001, 0.0181)
Southern	0.0074±0.0071 (0.0002, 0.0276)	**0.9948±0.0048 (0.9819, 0.9999)**

Values along the diagonal (bold) are the proportion of individuals derived from the source population (or non-migrant) each generation

The proportion of migrants received and given by each cluster was estimated using BAYESASS (table 2). It indicated low migration rates between clusters and very high migration rates within clusters, results that suggest a strong isolation between the two populations.

Fig. 4 STRUCTURE analysis tacking into account population origin of the samples as estimated by the Bayesian clustering methods and using the option MigrPrior = 0.05. Northern population represented by green and Southern by red

The analyses with STRUCTURE using the USEPOPINFO model and MigrPrior = 0.05 identified some admixed individuals and

migrants, but none presented a probability inferior to 0.5 of belonging to the originally assigned cluster (fig. 4). Bearing in mind the presence of null allele loci the cut-off value was lowered to 0.9, in this way when considering probabilities inferior to 0.9 (Appendix table 1) several individuals appear as admixed, which could be a sign of past gene flow

GENECLASS identified 10 migrants (Appendix table 1). Considering this result and the ones from STRUCTURE and BAPS concerning migrants, it was possible to detect eight migrants between the two clusters (represented by filled triangles in the distribution map) that were identified by at least two of the three methods.

Since the presence of null alleles could affect the evaluation of genetic variability in each cluster, I used the mean expected heterozygosity (He, Nei 1987) because the presence of null alleles has only a limited effect on this statistic (Chapuis & Estoup 2007). Overall, the 10 loci and both populations showed high genetic variability.

Correlation between genetic and geographic distances was assessed for each population by Mantel tests. For the Northern population it had a p-value of 0.043 and a Pearson correlation coefficient (R) of 0.063, while for the Southern population it had a p-value = 0.01 and R = 0.156. Although the null hypothesis of a random genetic distribution of individuals can be rejected, the support for isolation-by-distance (IBD) was low to moderate.

Table 3 Genetic diversity measures by populations for the ten microsatellite loci used. Allelic range (R), number of alleles (NA), allelic richeness (A_R), observed (H_O) and expected heterozygosity (H_E) and Hardy-Weinberg equilibrium P-value.

	Northern = 80						Southern = 123					
Locus	R	NA	A_R	H_O	H_E	P-value	R	NA	A_R	H_O	H_E	P-value
Sal 3	190-270	18	18	0.8	0.786	0.3623	180-254	22	20.040	0.797	0.890	**0.0005**
Sal E2	218-366	30	30	0.713	0.947	**0.0000**	226-366	36	33.940	0.846	0.963	**0.0000**
Sal E8	131-163	9	9	0.763	0.827	**0.0025**	131-171	11	10.837	0.813	0.874	0.0217
Sal E11	213-247	17	17	0.838	0.918	0.0062	209-259	23	21.110	0.805	0.868	**0.0030**
Sal E14	216-260	11	11	0.75	0.859	**0.0007**	216-272	15	13.894	0.512	0.846	**0.0000**
Sal E6	254-290	10	10	0.863	0.822	0.8897	250-302	12	11.285	0.772	0.793	0.1971
Sal E7	178-206	8	8	0.788	0.797	0.7068	178-214	9	8.286	0.675	0.756	0.012
SST-A6II	176-220	12	12	0.813	0.883	0.0077	172-228	14	13.518	0.886	0.893	0.6679
SST-C2	196-244	10	10	0.763	0.790	0.0269	204-264	13	11.546	0.732	0.783	0.1083
SST-B11	132-216	12	12	0.413	0.734	**0.0000**	132-212	17	15.298	0.715	0.748	**0.0020**

Loci that after Bonferroni correction departed significantly from Hardy–Weinberg equilibrium in each population are indicated in bold

The allele frequency distribution test showed that no shift in distribution could be detected for both populations, which remained in a normal L-shape (Appendix fig. 3), so the microsatellite data indicates that both populations did not undergo a recent bottleneck.

Mitochondrial DNA

The median-joining haplotype network based on cytochrome-*b* sequences of 64 samples and constructed with the software NETWORK revealed two main groups of haplotypes separated by 20 mutations (fig. 5A). The Northern population is mainly characterised by two divergent haplotypes separated by 10 mutations and one sample was identified as a migrant in the Southern population. In this population, four haplotypes were uncovered and a fifth closely-related haplotype was found in two individuals sampled in the Northern population.

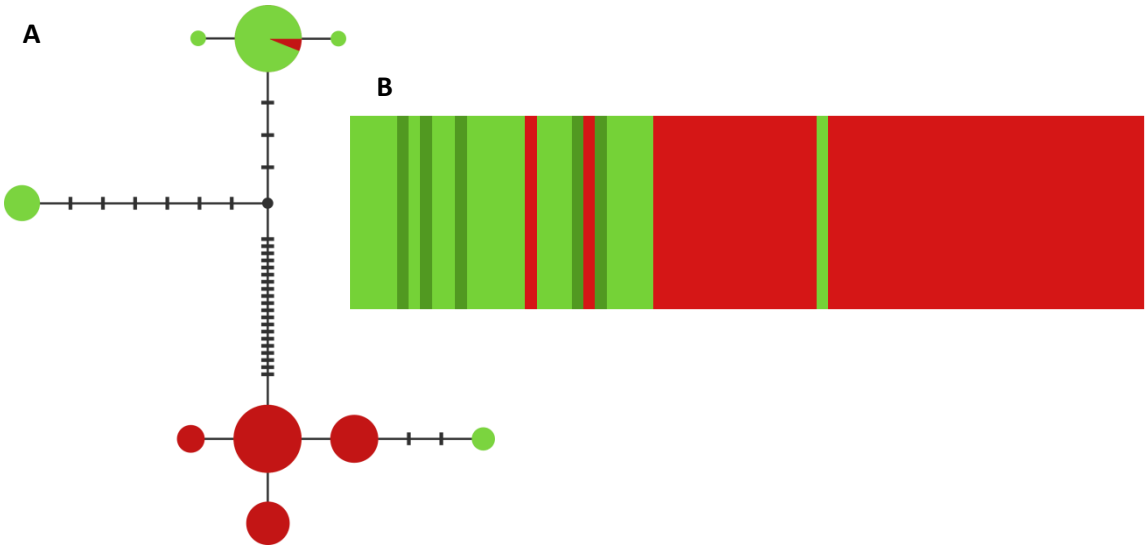


Fig. 5 mtDNA phylogeography and population structure of *Salamandra salamandra* in SW Portugal as inferred from cytochrome-*b* partial sequences of 64 individuals. Northern population represented by green and Southern by red. A: Median-joining haplotype network showing two groups of haplotypes separated by 20 mutations and a deep partition within the Northern population. B: Results from BAPS showing three clusters matching the three main haplogroups observed in the median-joining network.

BAPS produced similar results (fig. 5B) to the ones given by NETWORK, identifying 3 clusters: Southern and two clusters within the Northern population. It also detected the same three samples as from likely migrants. When comparing these migrant individuals detected from cytochrome-*b* data with the ones detected with microsatellites, only one individual was identified as a migrant by both markers. This individual is likely to represent a very recent migration event.

Estimates of genetic diversity obtained from the 702 bp cytochrome-*b* fragment are presented in table 4. Although more samples were analysed in the Southern population, with the exception of haplotype diversity, most diversity measures were lower in this population than in the Northern population. Mean population pairwise F_{ST} between Northern and Southern populations was high (0.75115).

Table 4 Genetic diversity estimates from a 702 bp cytochrome-*b* fragment. Number of individuals (N), haplotypes (N_H), nucleotide (π) and haplotype (h) diversity. The number of polymorphic sites is divided in transitions and transversions.

Populations	N	N_H	h	π	Polymorphic sites	
					Transitions	Transversions
Northern	26	5	0.5477	0.013137 ± 0.006932	31	3
Southern	38	4	0.6802	0.003094 ± 0.001951	24	2

The mismatch distribution computed for the Northern population showed, either using the sudden expansion model (P_{SSD} = 0.1659, P_{RAG} = 0.3246) or the spatial expansion model (P_{SSD} = 0.4777, P_{RAG} = 0.564), that this population is in demographic equilibrium. In the Southern population, the P_{SSD} = 0.0759 in the sudden expansion model also suggested demographic equilibrium but the P_{RAG} = 0.0145 of the same model and the spatial expansion model (P_{SSD} = 0.0077, P_{RAG} = 0.0132) indicate that this population should have undergone a recent demographic expansion.

Further assessment of past demography was provided by applying the neutrality tests Tajima's (1989) D and Fu's (1997) F_S. For the Northern population, the results of Tajima's D (D = 0.05967, P_D = 0.5767) were in agreement with a demographic equilibrium, whereas Fu's F_S (F_S = -23.49389, P_{FS} = 0.000) was not. For the southern population, past population expansion was supported by both tests (D = -2.29332, P_D = 0.0013; F_S = -26.93731, P_{FS} = 0.000).

The time since the beginning of this expansion was estimated using the mismatch distribution approach (Rogers & Harpending 1992) and the sudden and the spatial expansion models gave similar τ (tau): 1.053 and 1.052, respectively. This is respectively translated into an expansion time at 18,750 years before present (BP) with a 95% confidence interval of 4,274 – 32,835 years BP and at 18,732 years BP with a 95% confidence interval of 6,695 – 29,879 years BP.

The four simulations of the isolation with migration model implemented in IMa produced similar parameter estimates and distributions, with ESSs exceeding 200 for each estimated parameter. The results of IMa suggest that the Northern and Southern populations diverged 208,333 years BP (90% HPD = 103,785 – ?) considering a 0.8% per million years mutation rate and 5 years as generation time. The credibility intervals were recorded as the 90% HPD interval, which represents the shortest span that includes 90% of the probability density of the parameter, but the upper bound for the credibility interval is not given since it critically depends on the assumed prior for the maximum value of *t* when the curve slowly decreases to zero after the mode of *t* (Johnson *et al.* 2007; Guillaumet *et al.* 2008). The approach of Gaggiotti and Excoffier (2000) estimated the divergence between the Northern and Southern populations at 957,975 years BP (τ = 10.76). The population pairwise genetic distance D_A between

the two populations was estimated as 15.796 ($P<0.05$), which translates into a population divergence 1,406,339 years ago.

Discussion

The loci used in this work were chosen from the only two published panels of microsatellites for *Salamandra salamandra* (Steinfartz *et al.* 2004, Hendrix *et al.* 2010), but they presented some problems in the populations studied here in terms of null alleles. I tested other loci from the same panels but they were discarded because they showed weak amplification (Sal 29 and SST-G6) and binning problems likely due to false alleles (Sal23 and Sal E12). These difficulties were surprising because Hendrix *et al.* (2010) tested for the correlation between cross-amplification success and genetic distance within the Salamandridae family, and most of the used loci in this work had amplification success across species of the genus *Salamandra* (except for SST-C2 and Sal E11). This shows that although cross-amplification tests provide an useful preliminary insight of which primers might be employed in different species, there is no guarantee that the loci will be problem-free and not show deviations of Hardy-Weinberg equilibrium (HWE) caused by multiple distinct factors, even when marker transfer is just between subspecies of the species in which the microsatellites were isolated.

The determination of the population genetic structure in the dataset, previous to carrying out tests of HWE, linkage disequilibrium (LD) and null alleles, was important to remove the Wahlund effect (which was shown to be strong here) before assessing the quality of each locus for population genetic analysis. This approach allowed identifying the only loci (two) that showed unambiguous and consistent signs of null alleles.

ML-RELATE was used to detect samples with parent-offspring or full-sibling relationships, so that one of the individuals in each such pair could be removed. The use of this software allowed using the data of the two loci with null alleles because it has been shown (Wagner *et al.* 2006) that the method incorporated in ML-RELATE to deal with null alleles performs better than ignoring the presence of null alleles or discarding the affected loci.

Concerning the three Bayesian clustering methods, to check the impact of the null alleles, I repeated all the analyses without the two loci with null alleles identified in both clusters, and overall they gave the same results (Appendix fig. 1). To assess if the loci with null alleles could have an effect in measures of genetic distance and genetic differentiation I applied the correction method of Chapuis and Estoup (2007) implemented in FREENA. This approach indicated that the loci with null alleles did not influence significantly those measures and, therefore, final analyses were done using the 10 loci.

The three Bayesian clustering algorithms identified two clusters, one in the north of the study area that I designated "Northern" and can be associated with *S. s. gallaica* and the other that I designated "Southern" and can be associated with *S. s. crespoi*. Despite this congruent main result, slight differences were identified between the results of the three methods. GENELAND did not detect any migrants, BAPS showed some migrants, and STRUCTURE indicated some signs of admixture. Previous studies have noted that different clustering algorithms can infer different solutions. For instance, Rowe & Beebee (2007) reported non-congruent outputs from BAPS, GENELAND, and STRUCTURE, in terms of the estimated number of clusters. The results here do not show such level of incongruence between methods most likely because of the strong pattern of differentiation between the two subspecies, but it is worth noting that the slight differences detected could be due to differences in the underlying models, in the statistical estimators, or in the approximations in the algorithms used to compute the estimators. It is important to mention that all the programs use the genetic information to ascertain population membership of individuals without assuming predefined populations, operate by minimizing Hardy–Weinberg and linkage disequilibria, and the assignment of each individual genotype to its population of origin is carried out probabilistically using Markov chain Monte Carlo (MCMC). The use of MCMC is prone to convergence issues, this is, the software outputs might not be in certain cases the exact solution of the mathematical equations but an approximation which quality remains unknown.

The two obtained clusters seem separated by a narrow gap of 20Km in width that corresponds, approximately, to an area where the Sado River passes plus a belt of sandy soil south of it. Results from the analyses of the mitochondrial gene cytochrome-*b* indicated a similar partition boundary. Finally, despite repeated and intensive sampling efforts, no individuals could be found within the area. The fact that both mtDNA and microsatellites concur in suggesting a very reduced gene flow across the gap indicates that it has been a strong barrier both historically and contemporarily. Of all the methods used, only the ordination of individuals using the factorial correspondence analysis (FCA) did not show a discrete separation of the populations. However, it is known that the FCA may not have the needed resolution power when genetic differentiation as measured by F_{ST} is low, as is the case here (F_{ST} = 0.034). Despite strong isolation between populations, F_{ST} values can be low if gene flow, albeit very restricted, is historically recurrent.

The low migration rates detected with the software BAYESASS between the two subspecies are consistent with the hypothesis of a strong barrier separating them. Additionally, BAPS only identified three migrants, STRUCTURE (when defining the populations a priori) de-

tected few admixed individuals, and GENECLASS found 10 migrants. When considering as migrants only those individuals indicated by two or more software, eight individuals were identified among the 203 analysed individuals, again consistent with a scenario of reduced gene flow. The analyses of mtDNA found three migrants within the 64 surveyed individuals. Of the eight migrants detected with microsatellites only four were sequenced for their mtDNA, but one individual was identified has migrant by both genetic markers (Appendix table 1). The three migrants identified by microsatellites but not by mtDNA are recent but most likely less than the individual in which both markers concur. The very low number of migrants in terms of mtDNA is evidence for a historically very reduced gene flow.

The two migrants detected in the Northern population by their mtDNA presented a haplotype closely related but distinct of the haplotypes found in all individuals of the Southern population, which suggests a different source. I hypothesise that these migrants possibly originated from a different area, south of the Guadiana River, of the distribution of *S. s. crespoi*. To test this hypothesis it is necessary to study in the future the contact zone of the two Portuguese subspecies in southeast Portugal.

It was detected isolation-by-distance (IBD) patterns in both populations but relatively weak as measured by the low correlation coefficients in the Mantel tests. However, it was not detected any indication of substructure within each population (Appendix fig. 4), a finding that is consistent with the presence of IBD. Possibly, the IBD patterns detected in both populations are real but high correlation coefficients in the Mantel tests would only be observed with a more numerous and dense sampling from each population. Alternatively, the spatial arrangement of the genetic variation in the populations might be better described by an isolation-by-resistance (IBR) model which predicts a closer fit of genetic distance with resistance distance than with Euclidian geographic distance (McRae 2006). The IBR model attempts to account for the effect of habitat heterogeneity in shaping IBD patterns. To test this alternative, in addition to an improved sampling it would be necessary to construct a Geographic Information System (GIS) for the study area.

No signs of recent bottlenecks were detected in the microsatellite data from both populations. However, a bottleneck was suggested by the mtDNA data for the Southern population, which coherently exhibited the lowest nucleotide and haplotype diversities. Support for a sudden demographic expansion in the Southern population comes from its unimodal mismatch distribution and from a small value of the raggedness index (=0.0132). This suggests that, although both populations might be large for some time now, *S. s. crespoi* underwent

previously a expansion from a small population size. The time estimator (tau) of the mismatch distribution (MMD) dated this expansion around 18,000 years ago and this date is congruent between the two MMD models (sudden expansion and spatial expansion) used. This date is compatible with a scenario in which, during the Last Glacial Maximum (LGM) (between 26,500 and 19,000–20,000 years ago; Clark *et al.* 2009), *S. s. crespoi* suffered a severe demographic contraction and, after the LGM, the population recovered rapidly with the return of favourable climatic conditions. The support of the data for a spatial expansion of *S. s. crespoi* after the LGM opens the possibility that the present geographic proximity between the two subspecies in an almost parapatric distribution, might have been established relatively recently. This would be in agreement with the very reduced number of migrants detected between subspecies. Also, around 18,000 years ago almost all of the Portuguese continental shelf was above sea level, which allowed the confluence of northern Portuguese rivers (Rodrigues and Dias, 1989) and one may hypothesize that a similar phenomenon could have happened with the southern rivers. Enlarged and connected, the Sado and Guadiana Rivers, even if allowing some gene flow northwards, would have bogged down the expansion of *S. s. crespoi*. After the contraction and split of the river basins, the soil characteristics of many of the exposed riverbeds created strong barriers to dispersal. It is feasible that this scenario for the contraction and expansion of *S. s. crespoi* in southern Portugal has been replayed along the glacial-interglacial cycles of the Late Pleistocene.

The idea of this subspecies to have survived the LGM in a refuge in western Algarve, an area known to have been a refuge and a hotspot of evolutionary diversification for several *taxa* (Mesquita *et al.* 2005, Martinez-Solano *et al.* 2006, Fritz *et al.* 2006), would also agree with the fact that within this area (Serra de Monchique) the external morphological characteristics of the individuals are the more conspicuously typical of *S. s. crespoi*.

Three methods were used to estimate the divergence time between the two populations, as this parameter could provide further valuable insights into their evolutionary history. The estimate derived from D_A (Nei & Li 1979) was 1.4 million years and the method of Gaggiotti & Excoffier (2000) estimated it around 958,000 years ago. However, it needs to be emphasized that the two methods assume complete isolation of the daughter populations after the split, so they accommodate poorly instances of recurrent gene flow, as it seems to be the case here. In such cases, these methods tend to overestimate population divergence time and their results here fall somewhat halfway between the IMa estimate and the phylogenetic dating for the divergence of the lineages leading to the subspecies *crespoi* and *gallaica* (Steinfartz *et al.* 2000). The software IMa belongs to a recent set of programs, based in coalescent genealogies,

which generally are more powerful and robust than the two methods above to infer population parameters (Kuhner 2009). In particular, IMa is explicitly designed to take into account recurrent gene flow after population split (i.e. the last "moment" in which subpopulations were connected by "abundant" gene flow) (e.g. Palsboll *et al.* 2004). The IMa estimate was around 208,000 years BP.

Due to the stochastic nature of mutation and lineage sorting, estimated divergence times from genetic patterns must be considered only as rough approximations of the timing of historical events. Factors like the uncertainty on the substitution rate for the DNA region under study, unsampled populations, and small sample sizes, increase confidence limits on divergence time estimates. Keeping this in mind, I discuss below possible environmental correlations that could explain the indication by the mtDNA data of genetic connectivity between the two subspecies in the study area at around 210,000 years BP.

Around 220,000-210,000 years BP there is evidence for a period of warm and wet climate and expansion of *Quercus* forests in western Iberia (Desprat *et al.* 2006, Roucoux *et al.* 2006), which corresponds to a period of global favourable climate between the oxygen isotope stages (OIS) 7c and 7a. Besides a short peak of similar climate around 120,000 years ago, only today the climate in Europe is as favourable as at the OIS 7c-7a interval (Tzedakis *et al.* 1997). A warm and wet climate is likely to have led to an enlarged and tributary-rich Sado River that, together with an expanded forested habitat, would have facilitated gene flow to levels without any parallel since then. The fact that the Guadiana drainage was probably an affluent of the Sado River until the end of the Middle Pleistocene (Vidal *et al.* 1993) offers further support to the idea that expanded riverine connections, together with abundant forested land-cover, could have allowed substantial genetic exchange between the two subspecies across Alentejo.

It has been suggested that in amphibians, large rivers, such as Douro and Tagus, prevent dispersal of individuals between distant across river locations (Alexandrino *et al.* 2000; Martinez-Solano 2004). However, although the phylogeography of *S. s. gallaica* has not been examined yet in detail across its whole range, it is possible that this subspecies managed to cross these two large rivers. Moreover, although salamander adults seem to have a reduced dispersal capacity (Rebelo & Leclair 2003; Schmidt *et al.* 2007), juveniles or larvae can be transported by runoff water or floods) and undergo sweepstake dispersal (possibly across large distances).

The modelling of the distribution of *Salamandra salamandra* in south Portugal showed that a large percentage of the distribution can be explained by the absence of sand, pebbles

and clays (Appendix fig. 5; Simões *et al.* in prep.). This suggests that the strong barrier detected in this study is probably the consequence of the combined effect of the Sado River and the area of unsuitable habitat south of it. Presently, it is perhaps this belt of unsuitable habitat of sandy soil the most important component of the boundary between the two subspecies in southwest Portugal. The very reduced gene flow detected between these populations is interpreted as the result of rare and stochastic dispersal events up and down the Sado River.

In the near future I pretend to assess some of the hypotheses raised by this study, the first using a landscape genetics approach to examine the intraspecific patterns of the fire salamander in Portugal. Among them, the possible migration route across the Guadiana in the southeast of the country and, connected with this, the probable absence of the species in central Alentejo before it became an agricultural area in the Middle Ages. It will also be important to test the isolation-by-resistance hypothesis and correlate the resistance matrix obtained from an habitat suitability map with the genetic distance matrix, in order to better understand which landscape features and variables influence the genetic structure of the species.

References

Alcobendas M, Dopazo H, Alberch P (1996) Geographic variation in allozymes of populations of *Salamandra salamandra* (Amphibia: Urodela) exhibiting distinct reproductive modes. *Journal of Evolutionary Biology* 9:83-102.

Alexandrino J, Froufe E, Arntzen J, Ferrand N (2000) Genetic subdivision, glacial refugia and postglacial recolonization in the golden-striped salamander, *Chioglossa lusitanica* (Amphibia: Urodela). *Molecular Ecology* 9:771–781.

Almada VC, Pereira AM, Robalo JI, Fonseca JP, Levy A, Maia C, Valente A (2008) Mitochondrial DNA fails to reveal genetic structure in sea-lampreys along European shores. *Molecular Phylogenetics and Evolution* 46:391-6.

Almeida NF, Almeida PF, Gonçalves H, Sequeira F, Teixeira J, Almeida FF (2001) *Anfíbios e Répteis de Portugal*. Guias Fapas – Fundo para a Protecção dos Animais Selvagens. Porto: INOVA Artes Gráficas.

Bandelt H, Forster P, Rohl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* 16:37-48.

Barbujani G, Oden NL, Sokal RR (1989) Detecting Regions of Abrupt Change in Maps of Biological Variables. *Systematic Zoology* 38:376.

Bard E, Delaygue G, Rostek F, Antonioli F, Silenzi S, Schrag DP (2002) Hydrological conditions over the western Mediterranean basin during the deposition of the cold Sapropel 6 (ca. 175 kyr BP). *Earth and Planetary Science Letters* 202:481-494.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bohomme F (2004) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier, France.

Bernatchez L, Wilson CC (1998) Comparative phylogeography of Nearctic and Palearctic fishes. *Molecular Ecology* 7:431-452.

Bolfíková B, Hulva P (2011) Microevolution of sympatry: landscape genetics of hedgehogs *Erinaceus europaeus* and *E. roumanicus* in Central Europe. *Heredity*

Branco M, Ferrand N, Monnerot M (2000) Phylogeography of the European rabbit (*Oryctolagus cuniculus*) in the Iberian Peninsula inferred from RFLP analysis of the cytochrome-*b* gene. *Heredity* 85:307–317.

Branco M, Monnerot M, Ferrand N, Templeton AR (2002) Postglacial dispersal of the European rabbit (*Oryctolagus cuniculus*) on the Iberian Peninsula reconstructed from nested clade and mismatch analyses of mitochondrial DNA genetic variation. *Evolution* 56:792–803.

Brito RM, Briolay J, Galtier N, Bouvet Y, Coelho MM (1997) Phylogenetic relationships within genus *Leuciscus* (Pisces, Cyprinidae) in Portuguese fresh waters, based on mitochondrial DNA cytochrome *b* sequences. *Molecular Phylogenetics and Evolution* 8:435-42.

Caccone A, Milinkovitch M, Sbordoni V, Powell JR (1997) Mitochondrial DNA rates and biogeography in European newts (genus *Euproctus*). *Systematic Biology* 46:126.

Cavalli-Sforza LL, Edwards AW (1967) Phylogenetic analysis. Models and estimation procedures. *American Journal of Human Genetics* 19:233-57.

Centeno-Cuadros A, Delibes M, Godoy JA (2009) Phylogeography of Southern Water Vole (*Arvicola sapidus*): evidence for refugia within the Iberian glacial refugium? *Molecular Ecology* 18:3652-67.

Chapuis M-P, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution* 24:621-31.

Clark PU, Dyke AS, Shakun JD, Carlson AE, Clark J, Wohlfarth B, Mitrovica JX, Hostetler SW, McCabe AM (2009) The Last Glacial Maximum. *Science* 325:710-714.

Coelho M, Bogutskaya N, Rodrigues J, Collares-Pereira M (1998) *Leuciscus torgalensis*, and *L. aradensis*, two new cyprinids for Portuguese fresh waters. *Journal of Fish Biology* 52:937-950.

Coombs J, Letcher B, Nislow KH (2008) CREATE: a software to create input files from diploid genotypic data for 52 genetic software programs. *Molecular Ecology* 8:578-80.

Corander J, Waldmann P, Sillanpää MJ (2003) Bayesian analysis of genetic differentiation between populations. *Genetics* 163:367.

Corander J, Marttinen P (2006) Bayesian identification of admixture events using multilocus molecular markers. *Molecular Ecology* 15:2833-43.

Corander J, Tang J (2007) Bayesian analysis of population structure based on linked molecular information. *Mathematical Biosciences* 205:19-31.

Corander J, Marttinen P, Sirén J, Tang J (2008) Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. *BMC Bioinformatics* 9:539.

Coulon A, Guillot G, Cosson J-F, Angibault JM, Aulagnier S, Cargnelutti B, Galan M, Hewison JM (2006) Genetic structure is influenced by landscape features: empirical evidence from a roe deer population. *Molecular Ecology* 15:1669-79.

Cushman SA, McKelvey KS, Hayden J, Schwartz MK (2006) Gene flow in complex landscapes: testing multiple hypotheses with causal modeling. *The American Naturalist* 168:486-99.

Dempster AP, Laird NM, Rubin DB (1977) Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society B* 39:1-38.

Deter J, Chaval Y, Galan M, Gauffre B (2008) Kinship, dispersal and hantavirus transmission in bank and common voles. *Archives of Virology* 153:435-444.

Dieringer D, Schlötterer C (2003) Two distinct modes of microsatellite mutation processes: evidence from the complete genomic sequences of nine species. *Genome Research* 13:2242-51.

Dionne M, Caron F, Dodson JJ, Bernatchez L (2008) Landscape genetics and hierarchical genetic structure in Atlantic salmon: the interaction of gene flow and local adaptation. *Molecular Ecology* 17:2382-96.

Domínguez-Domínguez O, Boto L, Alda F, De León GPP, Doadrio I (2007) Human impacts on drainages of the Mesa Central, Mexico, and its genetic effects on an endangered fish, *Zoogoneticus quitzeoensis*. *Conservation Biology* 21:168-80.

Dupanloup I, Schneider S, Excoffier L (2002) A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology* 11:2571-2581.

Earl DA, VonHoldt BM (2011) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*.

Eiselt J (1958) Der Feuersalamander, Salamandra salamandra, Beiträge zu einer taxonomischen Synthese. *Abhandlungen und Berichte für Naturkunde und Vorgeschichte Museum Magdeburg* 10:77-154.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14:2611-20.

Excoffier L (2004) Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model. *Molecular Ecology* 13:853-864.

Fitzpatrick BM, Placyk JS, Niemiller ML, Casper GS, Burghardt GM (2008) Distinctiveness in the face of gene flow: hybridization between specialist and generalist gartersnakes. *Molecular Ecology* 17:4107-17.

Fitzpatrick BM, Shaffer HB (2007) Introduction history and habitat variation explain the landscape genetics of hybrid tiger salamanders. *Ecological Applications* 17:598–608.

Foley JE, Queen EV, Sacks B, Foley P (2005) GIS-facilitated spatial epidemiology of tick-borne diseases in coyotes (*Canis latrans*) in northern and coastal California. *Comparative Immunology, Microbiology and Infectious diseases* 28:197-212.

Fritz U, Barata M, Busack SD, Fritzsch G, Castilho R (2006) Impact of mountain chains, sea straits and peripheral populations on genetic and taxonomic structure of a freshwater turtle, *Mauremys leprosa* (Reptilia, Testudines, Geoemydidae). *Zoologica Scripta* 35:97-108.

Fu YX (1997) Statistical Tests of Neutrality of Mutations Against Population Growth, Hitchhiking and Background Selection. *Genetics* 147:915-925.

Funk WC, Blouin MS, Corn PS, Maxell BA, Pilliod DS, Amish S, Allendorf FW (2005) Population structure of Columbia spotted frogs (*Rana luteiventris*) is strongly affected by the landscape. *Molecular Ecology* 14:483-96.

Gaggiotti OE, Excoffier L (2000) A simple method of removing the effect of a bottleneck and unequal population sizes on pairwise genetic distances. *Proceedings of the Royal Society of London B* 267:81.

García-París M, Jockusch E (1999) A mitochondrial DNA perspective on the evolution of Iberian *Discoglossus* (Amphibia: Anura). *Journal of Zoology* 248:209–218.

Garcia-Paris M, Alcobendas M, Buckley D, Wake DB, García-París M (2003) Dispersal of viviparity across contact zones in Iberian populations of fire salamanders (*Salamandra*) inferred from discordance of genetic and morphological traits. *Evolution* 57:129–143.

Gasser F (1978) Le polytypisme de l'espece paléarctique *Salamandra salamandra* (L.) (Amphibien, Urodèle). II. Systèmatique, relations génétiques et tendences évolutives dans l'aire de répartition. *Archives De Zoologie Expérimentale et Générale* 119:635-668.

Goldberg CS, Waits LP (2010) Comparative landscape genetics of two pond-breeding amphibian species in a highly modified agricultural landscape. *Molecular Ecology* 19:3650-63.

Gómez A, Lunt DH (2007) Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. In Weiss S, Ferrand N (eds) *Phylogeography in southern European refugia: evolutionary perspectives on the origins and conservation of European biodiversity.* Kluwer, Dordrecht, pp 155–188.

Guicking D, Joger U, Wink M (2008) Molecular phylogeography of the viperine snake *Natrix maura* (Serpentes: Colubridae): evidence for strong intraspecific differentiation. *Organisms Diversity & Evolution* 8:130–145.

Guillaumet A, Crochet P-A, Pons J-M (2008) Climate-driven diversification in two widespread *Galerida* larks. *BMC Evolutionary Biology* 8:32.

Guillot G, Estoup A, Mortier F, Cosson JF (2005) A spatial statistical model for landscape genetics. *Genetics* 170:1261-80.

Guillot G, Mortier F, Estoup A (2005) GENELAND: a computer package for landscape genetics. *Molecular Ecology Notes* 5:712-715.

Guillot G (2008) Inference of structure in subdivided populations at low levels of genetic differentiation - the correlated allele frequencies model revisited. *Bioinformatics* 24:2222.

Guo SW, Thompson EA (1992) A Monte Carlo method for combined segregation and linkage analysis. *American Journal of Human Genetics* 51:1111-26.

Harpending HC (1994) Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology* 66:591-600.

Hendrix R, Hauswaldt JS, Veith M, Steinfartz S (2010) Strong correlation between cross-amplification success and genetic distance across all members of "True Salamanders" (Amphibia: Salamandridae) revealed by *Salamandra salamandra*-specific microsatellite loci. *Molecular Ecology Resources* 10:1038-1047.

Hewitt GM (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society* 68:87-112.

Hewitt GM (2001) Speciation, hybrid zones and phylogeography - or seeing genes in space and time. *Molecular Ecology* 10:537-49.

Hey J, Nielsen R (2007) Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proceedings of the National Academy of Sciences of the United States of America* 104:2785-90.

Holderegger R, Wagner H (2006) A brief guide to Landscape Genetics. *Landscape Ecology* 21:793-796.

Holderegger R, Wagner H (2008) Landscape genetics. *BioScience* 58:199–207.

Janssens X, Fontaine MC, Michaux JR, Libois R, De Kermabon J, Defourny P, Baret PV (2008) Genetic pattern of the recent recovery of European otters in southern France. *Ecography* 31:176–186.

Joger U, Steinfartz S (1994) Zur subspezifischen Gliederung der sudiberischen Feuersalamander (*Salamandra salamandra*, komplex). *Abhandl. Ber. Naturkde. Magdeburg* 17:83-98.

Johnson JA, Dunn PO, Bouzat JL (2007) Effects of recent population bottlenecks on reconstructing the demographic history of prairie-chickens. *Molecular Ecology* 16:2203-22.

Kalinowski ST, Wagner AP, Taper ML (2006) Ml-Relate: a Computer Program for Maximum Likelihood Estimation of Relatedness and Relationship. *Molecular Ecology Notes* 6:576-579.

Klewen R (1991) Die Landsalamander Europas, Teil 1: Die Gattungen *Salamandra* und *Mertensiella*. *Die Neue Brehm Bücherei* Nr.584.

Köhler G, Steinfartz S (2006) A new subspecies of the fire salamander, *Salamandra salamandra* (Linnaeus, 1758) from the Tendi valley, Asturias, Spain. *Salamandra* 42:13-20.

Kuhner MK (2009) Coalescent genealogy samplers: windows into population history. *Trends in Ecology and Evolution* 24:86-93.

Latch EK, Scognamillo DG, Fike JA, Chamberlain MJ, Rhodes OE (2008) Deciphering ecological barriers to North American river otter (*Lontra canadensis*) gene flow in the Louisiana landscape. *The Journal of Heredity* 99:265-74.

Leaché AD (2011) Multi-Locus Estimates of Population Structure and Migration in a Fence Lizard Hybrid Zone. *PloS one* 6:e25827.

Lecis R, Ferrando A, Ruiz-Olmo J, Mañas S, Domingo-Roura X, Manas S (2007) Population genetic structure and distribution of introduced American mink (*Mustela vison*) in Spain, based on microsatellite variation. *Conservation Genetics* 9:1149-1161.

Lessios HA, Garrido MJ, Kessing BD (2001) Demographic history of *Diadema antillarum*, a keystone herbivore on Caribbean reefs. *Proceedings of the Royal Society of London B* 268:2347-53.

Lisiecki LE, Raymo ME (2005) A Pliocene-Pleistocene stack of 57 globally distributed benthic $\delta^{18}O$ records. *Paleoceanography* 20:1-17.

Lowe WH, Likens GE, McPeek MA, Buso DC (2006) Linking Direct and Indirect Data on Dispersal: Isolation By Slope in a Headwater Stream Salamander. *Ecology* 87:334-339.

Maletzky A, Kaiser R, Mikulíček P (2010) Conservation Genetics of Crested Newt Species *Triturus cristatus* and *T. carnifex* within a Contact Zone in Central Europe: Impact of Interspecific Introgression and Gene Flow. *Diversity* 2:28-46.

Manel S, Berthoud F, Bellemain E, Gaudeul M, Luikart G, Swenson JE, Waits LP, Taberlet P (2007) A new individual-based spatial approach for identifying genetic discontinuities in natural populations. *Molecular Ecology* 16:2031-43.

Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology and Evolution* 18:189-197.

Manni F, Guerard E, Heyer E (2004) Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by using Monmonier's algorithm. *Human Biology* 76:173–190.

Martínez-Solano I, Alcobendas M, Buckley DM (2005) Molecular characterisation of the endangered *Salamandra salamandra almanzoris* (Caudata, Salamandridae). *Annales Zoologici*:57-68.

Martínez-Solano I, Gonçalves H, Arntzen J, García-París M (2004) Phylogenetic relationships and biogeography of midwife toads (Discoglossidae: *Alytes*). *Journal of Biogeography* 31:603–618.

Martínez-Solano I, González E (2008) Patterns of gene flow and source-sink dynamics in high altitude populations of the common toad *Bufo bufo* (Anura: Bufonidae). *Biological Journal of the Linnean Society* 95:824–839.

Martínez-Solano I, Teixeira J, Buckley DM, García-París M (2006) Mitochondrial DNA phylogeography of *Lissotriton boscai* (Caudata, Salamandridae): evidence for old, multiple refugia in an Iberian endemic. *Molecular Ecology* 15:3375-88.

Matschiner M, Hanel R, Salzburger W (2009) Gene flow by larval dispersal in the Antarctic notothenioid fish *Gobionotothen gibberifrons*. *Molecular Ecology* 18:2574-87.

Mesquita N, Hänfling B, Carvalho GR, Coelho MM (2005) Phylogeography of the cyprinid *Squalius aradensis* and implications for conservation of the endemic freshwater fauna of southern Portugal. *Molecular Ecology* 14:1939-54.

Miller MP, Haig SM, Wagner RS (2006) Phylogeography and spatial genetic structure of the Southern torrent salamander: implications for conservation and management. *The Journal of Heredity* 97:561-70.

Monmonier MS (1973) Maximum-Difference Barriers: An Alternative Numerical Regionalization Method. *Geographical Analysis* 5:245–261.

Mueller RL (2006) Evolutionary rates, divergence dates, and the performance of mitochondrial genes in Bayesian phylogenetic analysis. *Systematic Biology* 55:289-300.

Muñoz J, Gómez A, Green AJ, Figuerola J, Amat F, Rico C (2008) Phylogeography and local endemism of the native Mediterranean brine shrimp *Artemia salina* (Branchiopoda: Anostraca). *Molecular Ecology* 17:3160-77.

Murphy MA, Evans JS, Cushman SA, Storfer A (2008) Representing genetic variation as continuous surfaces: an approach for identifying spatial dependency in landscape genetic studies. *Ecography* 31:685-697.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences* 76:5269-5273.

Nielsen R, Wakeley J (2001) Distinguishing migration from isolation: a Markov chain Monte Carlo approach. *Genetics* 158:885-96.

Olalde M (2002) White oaks phylogeography in the Iberian Peninsula. *Forest Ecology and Management* 156:89-102.

Orsini L, Corander J, Alasentie A, Hanski I (2008) Genetic spatial structure in a butterfly metapopulation correlates better with past than present demographic structure. *Molecular ecology* 17:2629-42.

Paetkau D, Slade R, Burden M, Estoup A (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Molecular Ecology* 13:55-65.

Palsbøll PJ, Bérubé M, Aguilar A, Notarbartolo-Di-Sciara G, Nielsen R (2004) Discerning between recurrent gene flow and recent divergence under a finite-site mutation model applied to North Atlantic and Mediterranean Sea fin whale (*Balaenoptera physalus*) populations. *Evolution* 58(3): 670-675.

Paulo O, Dias C, Bruford M (2001) The persistence of Pliocene populations through the Pleistocene climatic cycles: evidence from the phylogeography of an Iberian lizard. *Proceedings of the Royal Society of London B* 268:1625-1630.

Paulo OS, Jordan WC, Bruford MW, Nichols RA (2002) Using nested clade analysis to assess the history of colonization and the persistence of populations of an Iberian Lizard. *Molecular Ecology* 11:809-19.

Peakall R, Smouse PE (2006) Genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6:288-295.

Piry S, Luikart G, Cornuet J-M (1999) BOTTLENECK: a computer program for detecting recent reductions in the effective size using allele frequency data. *Journal of Heredity* 90:502–503.

Piry S, Alapetite A, Cornuet J-M, Paetkau D, Baudouin L, Estoup A (2004) GeneClass2: A Software for Genetic Assignment and First-Generation Migrant Detection. *Journal of Heredity* 95:536-539.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of Population Structure Using Multilocus Genotype Data. *Genetics* 155:945-959.

Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences* 94:9197-9201.

Ray N, Currat M, Excoffier L (2003) Intra-deme molecular diversity in spatially expanding populations. *Molecular Biology and Evolution* 20:76.

Rebelo R, Caetano MH (1995) Use of the skeletochronolgical method for ecodemographical studies on *Salamandra salamandra gallaica* from Portugal. *In* Llorente G, Montori A, Santos X, Carretero MA (eds), *Scientia Herpetologica*, Eds, Barcelona, Asociación Herpetológica Española, pp 135-140.

Rebelo R (2002) Biologia e ecodemografia comparadas de *Salamandra salamandra* (L., 1758) em Portugal. Ph.D. diss., Universidade of Lisbon, Lisbon, Portugal.

Rebelo R, Leclair MH (2003) Site Tenacity in the Terrestrial Salamandrid *Salamandra salamandra*. *Journal of Herpetology* 37:440-445.

Rebelo R (2008) *Salamandra salamandra*. *In* Loureiro A, Ferrand de Almeida N, Carretero M, Paulo O (eds) *Atlas dos Anfíbios e Répteis de Portugal*, Instituto da Conservação da Natureza e da Biodiversidade, Lisboa, pp 96-97.

Rice WR (1989) Analyzing tables of statistical tests. *Evolution* 43:223–225.

Rodrigues A, Dias JMA (1989) Evolução pós-glaciária da plataforma continental Portuguesa a norte do cabo Mondego. *Anais Instituto Hidrográfico* 10:39-50.

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution* 9:552-69.

Roucoux KH, Tzedakis PC, de Abreu L, Shackleton NJ (2006) Climate and vegetation changes 180,000 to 345,000 years ago recorded in a deep-sea core off Portugal. *Earth and Planetary Science Letters* 249:307-325.

Rousset F (2008) Genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources* 8:103-6.

Rowe G, Beebee TJC (2007) Defining population boundaries: use of three Bayesian approaches with microsatellite data from British natterjack toads (*Bufo calamita*). *Molecular ecology* 16:785-96.

Safner T, Miaud C, Gaggiotti O, Decout S, Rioux D, Zundel S, Manel S (2010) Combining demography and genetic analysis to assess the population structure of an amphibian in a human-dominated landscape. *Conservation Genetics* 12:161-173.

Safner T, Miller MP, McRae BH, Fortin M-J, Manel S (2011) Comparison of Bayesian Clustering and Edge Detection Methods for Inferring Boundaries in Landscape Genetics. *International Journal of Molecular Sciences* 12:865-889.

Salvador A (1974) *Guía de los anfibios y reptiles españoles*, ICONA, Madrid.

Salvador A, García-París M (2001) *Anfibios Españoles. Identificación, Historia Natural y Distribución*, Canseco Editores, S. L., Talavera de la Reina.

Schmidt BR, Schaub M, Steinfartz S (2007) Apparent survival of the salamander *Salamandra salamandra* is low because of high migratory activity. *Frontiers in zoology* 4:19.

Schneider S, Excoffier L (1999) Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics* 152:1079-89.

Segelbacher G, Manel S, Tomiuk J (2008) Temporal and spatial analyses disclose consequences of habitat fragmentation on the genetic diversity in capercaillie (*Tetrao urogallus*). *Molecular Ecology* 17:2356-67.

Seutin G, White BN, Boag PT (1991) Preservation of avian blood and tissue samples for DNA analyses. *Canadian Journal of Zoology* 69:82-90.

Slatkin M, Hudson RR (1991) Pairwise Comparisons of Mitochondrial DNA Sequences in Stable and Exponentially Growing Populations. *Genetics* 129:555-562.

Spear SF, Storfer A (2008) Landscape genetic structure of coastal tailed frogs (*Ascaphus truei*) in protected vs. managed forests. *Molecular Ecology* 17:4642-56.

Steinfartz S, Kusters D, Tautz D (2004) Isolation and characterization of polymorphic tetranucleotide microsatellite loci in the Fire salamander *Salamandra salamandra* (Amphibia: Caudata). *Molecular Ecology Notes* 4:626-628.

Steinfartz S, Veith M, Tautz D (2000) Mitochondrial sequence analysis of *Salamandra* taxa suggests old splits of major lineages and postglacial recolonizations of Central Europe from distinct source populations of *Salamandra salamandra*. *Molecular Ecology* 9:397–410.

Storfer A, Murphy MA, Evans JS, Goldberg CS, Robinson S, Spear SF, Dezzani R, Delmelle E, Vierling L, Waits LP (2007) Putting the "landscape" in landscape genetics. *Heredity* 98:128-42.

Storfer A, Murphy MA, Spear SF, Holderegger R, Waits LP (2010) Landscape genetics: where are we now? *Molecular Ecology*:3496-3514.

Tajima F (1989) Statistical Method for Testing the Neutral Mutation Hypothesis by DNA Polymorphism. *Genetics* 123:585-595.

Tajima F (1989) The Effect of Change in Population Size on DNA Polymorphism. *Genetics* 123:597-601.

Tajima F (1993) Measurement of DNA polymorphism. *In* Takahata N and Clark AG (eds) *Mechanisms of Molecular Evolution. Introduction to Molecular Paleopopulation Biology*, MA:Japan Scientific Societies Press, Sinauer Associates, Inc., Sunderland, MA, Tokyo, pp 37-59.

Tan A-M, Wake DB (1995) MtDNA Phylogeography of the California Newt, *Taricha torosa* (Caudata, Salamandridae). *Molecular Phylogenetics and Evolution* 4:383-94.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) Micro-Checker: Software for Identifying and Correcting Genotyping Errors in Microsatellite Data. *Molecular Ecology Notes* 4:535-538.

Veith M (1994) Morphological, molecular and life history variation in *Salamandra salamandra* (L.). *Mertensiella* 4:355–398.

Vialatte A, Guiller A, Bellido A, Madec L (2008) Phylogeography and historical demography of the Lusitanian snail *Elona quimperiana* reveal survival in unexpected separate glacial refugia. *BMC Evolutionary Biology* 8:339.

Vidal JR, Cáceres LM, Ramirez AR (1993) Modelo evolutivo de la red fluvial Cuaternaria en el suroeste de la Peninsula Iberica. *In Actas da 3ª Reunião do Quaternário Ibérico*, Universidade de Coimbra, pp 93-96.

Wagner AP, Creel S, Kalinowski ST (2006) Estimating relatedness and relationships using microsatellite loci with null alleles. *Heredity* 97:336-45.

Wagner HH, Werth S, Kalwij JM, Bolli JC, Scheidegger C (2006) Modelling forest recolonization by an epiphytic lichen using a landscape genetic approach. *Landscape Ecology* 21:849-865.

Wang Y-H, Yang K-C, Bridgman CL, Lin L-K (2008) Habitat suitability modelling to correlate gene flow with landscape connectivity. *Landscape Ecology*:989-1000.

Waples RS, Gaggiotti O (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology* 15:1419-39.

Weir BS (1996) *Genetic Data Analysis II*, Sinauer Associates, Inc., Sunderland, MA.

Wilson GA, Rannala B (2003) Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* 163:1177-91.

Womble W (1951) Differential systematics. *Science* 114:315-322.

Wood MJ, Cosgrove CL, Wilkin TA, Knowles SCL, Day KP, Sheldon BC (2007) Within-population variation in prevalence and lineage distribution of avian malaria in blue tits, *Cyanistes caeruleus*. *Molecular Ecology* 16:3263-73.

Wilmer JW, Elkin C, Wilcox C, Murray L, Niejalke D, Possingham H (2008) The influence of multiple dispersal mechanisms and landscape structure on population clustering and connectivity in fragmented artesian spring snail populations. *Molecular Ecology* 17:3733-51.

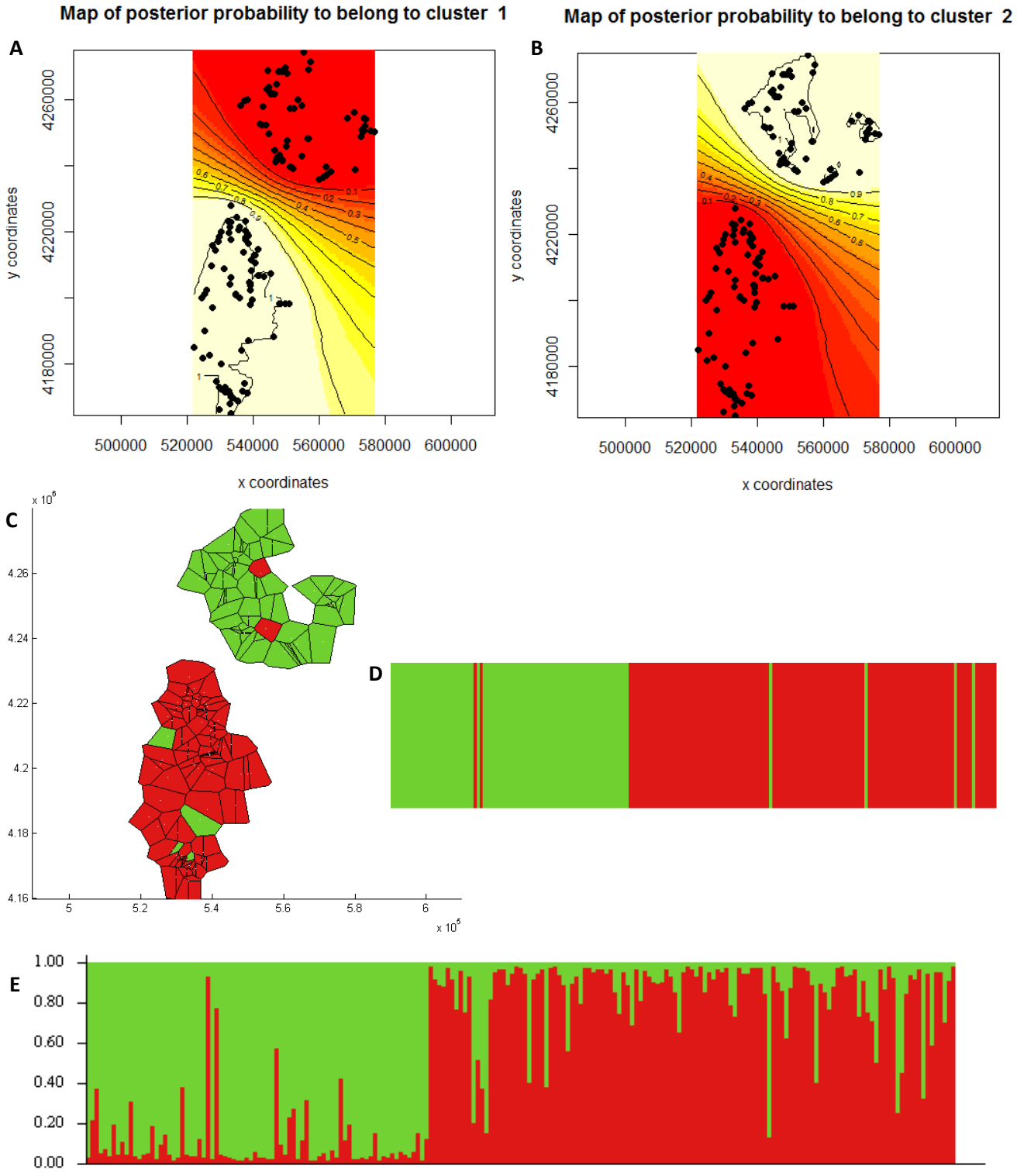
Wright S (1931) Evolution in Mendelian Populations. *Genetics* 16:97-159.

Zamudio KR, Savage WK (2003) Historical isolation, range expansion, and secondary contact of two highly divergent mitochondrial lineages in spotted salamanders (*Ambystoma maculatum*). *Evolution* 57:1631–1652.

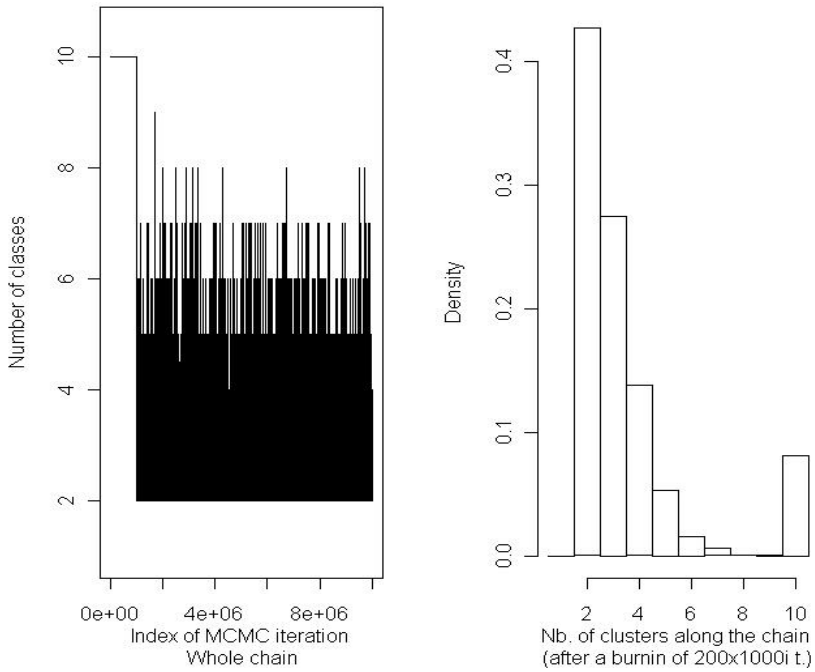
Zannèse A, Morellet N, Targhetta C, Coulon A, Fuser S, Hewison AJM, Ramanzin M (2006) Spatial structure of roe deer populations: towards defining management units at a landscape scale. *Journal of Applied Ecology* 43:1087-1097.

Zardoya R, Doadrio I (1998) Phylogenetic relationships of Iberian cyprinids: systematic and biogeographical implications. *Proceedings of the Royal Society of London B* 265:1365-72.

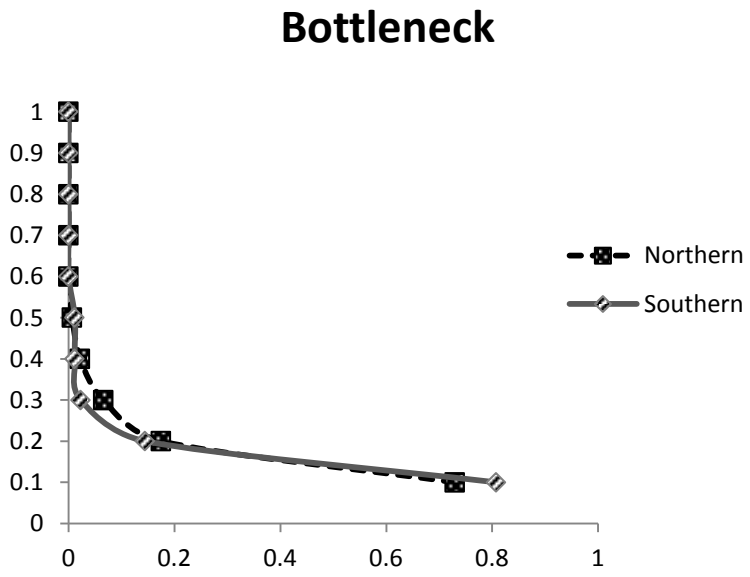
Appendix



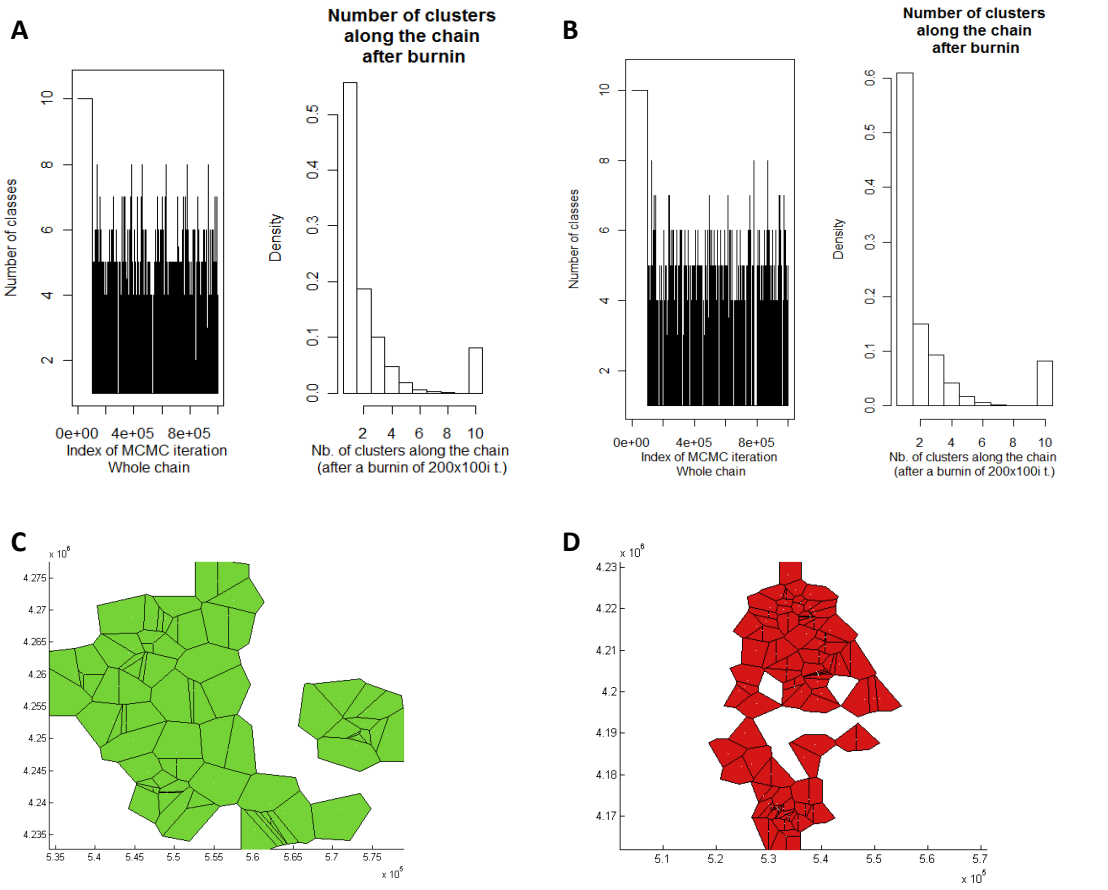
Appendix fig. 1 Bayesian clustering results based on the 8 microsatellite loci without null alleles. A and B: maps of posterior probability for each cluster accordingly with GENELAND. These maps are a tessellations analysis representation with the probability that a sample belongs to a particular group ranging from low (dark colour) to high (light colour). C and D: results from BAPS; C is a spatial Voronoi tessellation representation of the admixture plot (D) in which each bar represents an individual. Northern population represented by green and Southern by red. Another two migrants were detected without the two null alleles loci. E: STRUCTURE plot for the best-supported *k* (=2) as inferred with the method of Evanno *et al.* (2005); each bar represents an individual. Northern population represented by green and Southern by red.



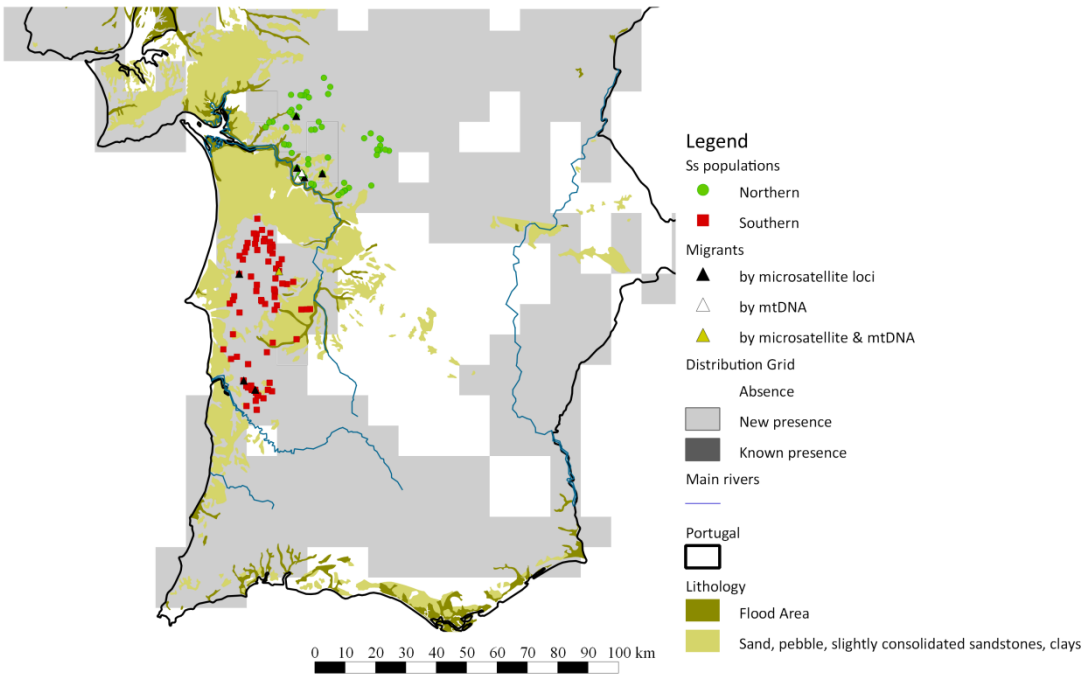
Appendix fig. 2 GENELAND output result that shows the number of clusters obtained along the chain after burnin.



Appendix fig. 3 BOTTLENECK results: Allele frequency distribution in a normal L-shape for both populations indicating that neigher has undergone a recent bottleneck



Appendix fig. 4 Bayesian clustering results based on the 10 microsatellite loci for Northern (A and C) and Southern (B and D) populations. A and B: GENELAND output result that shows no substructure within each population. C and D: BAPS spatial Voronoi tessellation representation of the admixture plot confirming no substructure within each population.



Appendix fig. 5 *Salamandra salamandra* sampling locations and sandy soil distribution. Salamander known distribution (Grey) according to Rebelo (2008).

Appendix table 1 Samples (N), UTM geographic coordinates (X, Y), cluster assigned to each sample by the Bayesian clustering methods and the three programs used to identify the migrants. BAPS population assignment to cluster 1 or to cluster 2, to GENECLASS is presented the -$\log_{10}$ (L_home /L_max) and nonmigrant probability (p-value<0.05); and to STRUCTURE is presented q-values when using a MIGRPRIOR<0.05. Within each analysis significant values are indicated with dark grey (in STRUCTURE the cut-off is 0.9), migrants identified by two or more programs are indicated with light grey. In bold, samples also analysed by mitochondrial DNA. * Migrants only detected by the 8 loci analysis. ** Migrants obtained by mitochondrial DNA analyses.

N	X	Y	Cluster assigned	BAPS		GENECLASS		STRUCTURE	
				1	2	-$\log_{10}$(L_home /L_max)	Prob.	1	2
1	576876	4250488	Northern	1	0	0	0.5107	0.985	0.015
2	575598	4250837	Northern	1	0	0	0.511	0.948	0.052
3	574188	4254169	Northern	1	0	0	0.5107	0.954	0.046
4	**573983**	**4252153**	**Northern**	**1**	**0**	**0**	**0.5113**	**0.97**	**0.03**
5	573771	4254540	Northern	1	0	0	0.5108	0.985	0.015
6	573725	4250676	Northern	1	0	0	0.51	0.97	0.03
7	573243	4250342	Northern	1	0	0	0.5115	0.92	0.08
8	572981	4250882	Northern	1	0	0	0.5113	0.987	0.013
9	572969	4250901	Northern	1	0	0	0.5117	0.961	0.039
10	**572841**	**4249067**	**Northern**	**1**	**0**	**0**	**0.5106**	**0.924**	**0.076**
11	**571035**	**4239030**	**Northern**	**1**	**0**	**0.206**	**0.0194**	**0.965**	**0.035**
12	**570841**	**4256215**	**Northern**	**1**	**0**	**0**	**0.51**	**0.975**	**0.025**
13	568678	4254599	Northern	1	0	0	0.5111	0.994	0.006
14	563689	4238268	Northern	1	0	0	0.512	0.986	0.014
15	**562414**	**4237414**	**Northern**	**1**	**0**	**0**	**0.5122**	**0.989**	**0.011**
16	562148	4237275	Northern	1	0	0	0.5115	0.939	0.061
17	562147	4237275	Northern	1	0	0	0.5117	0.965	0.035
18	**562085**	**4239943**	**Northern**	**1**	**0**	**0**	**0.5119**	**0.953**	**0.047**
19	561520	4236840	Northern	1	0	0	0.511	0.978	0.022
20	**561326**	**4236699**	**Northern**	**1**	**0**	**0**	**0.5109**	**0.984**	**0.016**
21	560206	4235872	Northern	1	0	0	0.5102	0.994	0.006
22	**557511**	**4271515**	**Northern**	**1**	**0**	**0**	**0.5104**	**0.991**	**0.009**
23	556819	4269161	Northern	1	0	0	0.5092	0.941	0.059
24	556818	4269161	Northern	1	0	0	0.5098	0.967	0.033
25	556755	4248257	Northern	1	0	0	0.5105	0.988	0.012
26	**556710**	**4248264**	**Northern**	**1**	**0**	**0**	**0.5096**	**0.991**	**0.009**
27	555419	4274491	Northern	1	0	0	0.5095	0.926	0.074
28	555418	4274491	Northern	1	0	0	0.5108	0.983	0.017
29	**554900**	**4243022**	**Northern**	**0**	**1**	**2.59**	**0.0013**	**0.596**	**0.404**
30	**554713**	**4258308**	**Northern**	**1**	**0**	**0**	**0.5113**	**0.992**	**0.008**
31*	553721	4260027	Northern	1	0	1.11	0.0062	0.908	0.092
32	552442	4257462	Northern	1	0	0	0.5105	0.989	0.011
33	552441	4257462	Northern	1	0	0	0.5117	0.981	0.019
34	**552084**	**4239087**	**Northern**	**1**	**0**	**0**	**0.5107**	**0.987**	**0.013**
35	552017	4239402	Northern	1	0	0	0.5119	0.982	0.018
36	551236	4257377	Northern	1	0	0	0.5105	0.995	0.005
37	551221	4239679	Northern	1	0	0	0.5107	0.994	0.006
38	550458	4268039	Northern	1	0	0	0.5096	0.99	0.01
39	550457	4268039	Northern	1	0	0	0.5115	0.993	0.007
40	**550279**	**4247768**	**Northern**	**1**	**0**	**0**	**0.5101**	**0.976**	**0.024**
41	550272	4246020	Northern	1	0	0	0.5115	0.985	0.015
42	550271	4246020	Northern	1	0	0	0.5122	0.988	0.012
43	549735	4269811	Northern	1	0	0	0.5113	0.995	0.005
44	549014	4241697	Northern	1	0	0	0.5098	0.992	0.008
45	**549013**	**4241737**	**Northern**	**1**	**0**	**0.02**	**0.0193**	**0.733**	**0.267**
46	549013	4241697	Northern	1	0	0	0.5114	0.88	0.12
47	549012	4241697	Northern	1	0	0	0.51	0.99	0.01
48	549007	4241509	Northern	1	0	0	0.5099	0.886	0.114
49	**548867**	**4268714**	**Northern**	**1**	**0**	**0**	**0.5108**	**0.976**	**0.024**
50	548178	4268738	Northern	1	0	0	0.5107	0.993	0.007
51**	**548134**	**4243237**	**Northern**	**1**	**0**	**0**	**0.5105**	**0.956**	**0.044**
52	547683	4268772	Northern	1	0	0	0.5104	0.916	0.084
53	**547359**	**4242650**	**Northern**	**1**	**0**	**0**	**0.5126**	**0.996**	**0.004**
54	547276	4242314	Northern	1	0	0	0.5113	0.987	0.013
55	547275	4242314	Northern	1	0	0	0.5107	0.985	0.015
56	547272	4264820	Northern	1	0	0	0.5102	0.991	0.009
57	547271	4264821	Northern	1	0	0	0.5108	0.997	0.003
58	547271	4264820	Northern	1	0	0	0.51	0.938	0.062
59**	**546793**	**4241353**	**Northern**	**1**	**0**	**0**	**0.5102**	**0.991**	**0.009**
60	546580	4244890	Northern	1	0	0.344	0.0143	0.863	0.137
61	**546465**	**4261883**	**Northern**	**1**	**0**	**0**	**0.5093**	**0.958**	**0.042**
62	546311	4261900	Northern	1	0	0.553	0.0126	0.879	0.121
63	**545461**	**4262020**	**Northern**	**1**	**0**	**0**	**0.5104**	**0.992**	**0.008**
64	545202	4262332	Northern	1	0	0	0.5115	0.993	0.007
65	**544829**	**4249768**	**Northern**	**1**	**0**	**0**	**0.511**	**0.984**	**0.016**
66	**544756**	**4263971**	**Northern**	**1**	**0**	**0**	**0.5126**	**0.995**	**0.005**

N	X	Y	Cluster assigned	BAPS		GENECLASS		STRUCTURE	
				1	2	$-\log_{10}$(L_home /L_max)	Prob.	1	2
67	**544608**	**4268901**	**Northern**	**1**	**0**	**0**	**0.5102**	**0.98**	**0.02**
68	544211	4263245	Northern	1	0	0	0.5117	0.977	0.023
69	543879	4252422	Northern	1	0	0	0.511	0.996	0.004
70	543878	4252422	Northern	1	0	0	0.5108	0.979	0.021
71	**542984**	**4258171**	**Northern**	**1**	**0**	**0**	**0.5112**	**0.994**	**0.006**
72	542983	4258171	Northern	1	0	0	0.5108	0.982	0.018
73	542983	4258171	Northern	1	0	0	0.5118	0.984	0.016
74	542725	4252479	Northern	1	0	0	0.5111	0.979	0.021
75	**542274**	**4252627**	**Northern**	**1**	**0**	**0**	**0.5101**	**0.984**	**0.016**
76	538490	4260018	Northern	1	0	0	0.5112	0.984	0.016
77	538490	4260019	Northern	1	0	0	0.5119	0.985	0.015
78	**538489**	**4260018**	**Northern**	**1**	**0**	**0**	**0.5104**	**0.984**	**0.016**
79	537425	4259859	Northern	1	0	0	0.5112	0.994	0.006
80	536253	4258251	Northern	1	0	0	0.5106	0.986	0.014
81	**550867**	**4198363**	**Southern**	**0**	**1**	**0**	**0.5066**	**0.003**	**0.997**
82	549844	4198286	Southern	0	1	0	0.5069	0.015	0.985
83	547987	4198223	Southern	0	1	0	0.5062	0.011	0.989
84	**546429**	**4188437**	**Southern**	**0**	**1**	**0**	**0.5065**	**0.024**	**0.976**
85	546428	4188437	Southern	0	1	0	0.5074	0.01	0.99
86	545329	4207330	Southern	0	1	0	0.5069	0.023	0.977
87	545328	4207330	Southern	0	1	0	0.5056	0.129	0.871
88	543265	4206592	Southern	0	1	0	0.5067	0.012	0.988
89	541841	4206750	Southern	0	1	0	0.5067	0.035	0.965
90	541727	4206775	Southern	0	1	0	0.5066	0.026	0.974
91	**541539**	**4214768**	**Southern**	**0**	**1**	**0**	**0.5065**	**0.041**	**0.959**
92	540691	4213149	Southern	0	1	0	0.5067	0.04	0.96
93	540690	4213149	Southern	0	1	0	0.5068	0.08	0.92
94**	**540633**	**4210828**	**Southern**	**0**	**1**	**2.695**	**0.0003**	**0.294**	**0.706**
95	540632	4210828	Southern	0	1	0	0.5065	0.024	0.976
96	540507	4213150	Southern	0	1	0	0.5068	0.056	0.944
97	**539949**	**4199591**	**Southern**	**0**	**1**	**0**	**0.5063**	**0.016**	**0.984**
98	539671	4211709	Southern	0	1	0	0.5062	0.006	0.994
99	539470	4208444	Southern	0	1	0	0.5071	0.032	0.968
100	539469	4208444	Southern	0	1	0	0.5069	0.013	0.987
101	539308	4203941	Southern	0	1	0	0.5069	0.004	0.996
102	539299	4204080	Southern	0	1	0	0.5065	0.009	0.991
103	539235	4204319	Southern	0	1	0	0.5061	0.015	0.985
104	539207	4202332	Southern	0	1	0	0.5063	0.056	0.944

N	X	Y	Cluster assigned	BAPS		GENECLASS		STRUCTURE	
				1	2	$-\log_{10}$(L_home /L_max)	Prob.	1	2
105	539194	4198118	Southern	0	1	0	0.5072	0.048	0.952
106	**539164**	**4204540**	**Southern**	**0**	**1**	**0**	**0.5067**	**0.014**	**0.986**
107	539011	4204925	Southern	0	1	0	0.5063	0.016	0.984
108	**538588**	**4218911**	**Southern**	**0**	**1**	**0**	**0.5057**	**0.346**	**0.654**
109	538583	4187325	Southern	0	1	0	0.5067	0.009	0.991
110	538529	4216705	Southern	0	1	0	0.5065	0.003	0.997
111	538528	4216705	Southern	0	1	0	0.507	0.02	0.98
112	538368	4219621	Southern	0	1	0	0.5069	0.018	0.982
113	538344	4171314	Southern	0	1	0	0.5071	0.075	0.925
114	538343	4171314	Southern	0	1	0	0.5068	0.044	0.956
115	537798	4217624	Southern	0	1	0	0.506	0.029	0.971
116	**537796**	**4223384**	**Southern**	**0**	**1**	**0**	**0.5073**	**0.031**	**0.969**
117	537705	4217974	Southern	0	1	0	0.5073	0.005	0.995
118	537704	4217975	Southern	0	1	0	0.5068	0.003	0.997
119	537704	4217974	Southern	0	1	0	0.5066	0.014	0.986
120	537702	4218464	Southern	0	1	0	0.5069	0.01	0.99
121	537701	4218464	Southern	0	1	0	0.5071	0.022	0.978
122	537634	4220496	Southern	0	1	0	0.5067	0.014	0.986
123	**537546**	**4174109**	**Southern**	**0**	**1**	**0**	**0.5063**	**0.015**	**0.985**
124	537545	4174109	Southern	0	1	0	0.5067	0.043	0.957
125	**537137**	**4214027**	**Southern**	**0**	**1**	**0**	**0.5074**	**0.172**	**0.828**
126	536982	4171768	Southern	0	1	0	0.5071	0.026	0.974
127	536981	4171768	Southern	0	1	0	0.5071	0.063	0.937
128	**536654**	**4184158**	**Southern**	**0**	**1**	**0**	**0.5068**	**0.044**	**0.956**
129	536066	4200068	Southern	0	1	0	0.5062	0.015	0.985
130	**535932**	**4221498**	**Southern**	**0**	**1**	**0**	**0.5063**	**0.045**	**0.955**
131	**535930**	**4221500**	**Southern**	**0**	**1**	**0**	**0.5065**	**0.017**	**0.983**
132	**535608**	**4220650**	**Southern**	**0**	**1**	**0**	**0.5069**	**0.019**	**0.981**
133	**535543**	**4169000**	**Southern**	**0**	**1**	**0**	**0.5068**	**0.02**	**0.98**
134	535542	4169000	Southern	0	1	0	0.5063	0.124	0.876
135	**535047**	**4224424**	**Southern**	**0**	**1**	**0**	**0.5065**	**0.016**	**0.984**
136	**534922**	**4200961**	**Southern**	**0**	**1**	**0**	**0.5059**	**0.004**	**0.996**
137	534770	4201355	Southern	0	1	0	0.5065	0.016	0.984
138	534769	4201355	Southern	0	1	0	0.5074	0.016	0.984
139	533812	4169830	Southern	0	1	0	0.5069	0.021	0.979
140	533811	4169830	Southern	0	1	0	0.5069	0.006	0.994
141	533510	4170071	Southern	0	1	0	0.5071	0.003	0.997
142	533509	4170071	Southern	0	1	0	0.5067	0.008	0.992

N	X	Y	Cluster assigned	BAPS		GENECLASS		STRUCTURE	
				1	2	$-\log_{10}$(L_home /L_max)	Prob.	1	2
143	**533503**	**4228104**	**Southern**	**0**	**1**	**0**	**0.5074**	**0.013**	**0.987**
144	**533495**	**4217816**	**Southern**	**0**	**1**	**0**	**0.5068**	**0.017**	**0.983**
145	533494	4217816	Southern	0	1	0	0.5065	0.043	0.957
146	533348	4206347	Southern	0	1	0	0.5067	0.015	0.985
147	533347	4206347	Southern	0	1	0	0.5059	0.014	0.986
148	533310	4165221	Southern	0	1	0	0.5073	0.013	0.987
149	533310	4165220	Southern	0	1	0	0.5067	0.014	0.986
150	533309	4165220	Southern	0	1	0	0.5073	0.008	0.992
151	533284	4223111	Southern	0	1	0	0.5072	0.018	0.982
152	533283	4223111	Southern	0	1	0	0.5065	0.043	0.957
153	533197	4204305	Southern	0	1	0	0.5065	0.009	0.991
154	533196	4204305	Southern	0	1	0	0.506	0.023	0.977
155	533097	4221505	Southern	0	1	0	0.5057	0.009	0.991
156	533039	4170868	Southern	0	1	0	0.5072	0.007	0.993
157	533038	4170868	Southern	0	1	0	0.5069	0.013	0.987
158	532998	4168179	Southern	0	1	0	0.5076	0.005	0.995
159	532997	4168179	Southern	0	1	0	0.506	0.034	0.966
160*	532822	4171800	Southern	0	1	0.09	0.0114	0.301	0.699
161	532821	4171800	Southern	0	1	0	0.507	0.003	0.997
162	532776	4219839	Southern	0	1	0	0.5063	0.03	0.97
163	**532344**	**4223287**	**Southern**	**0**	**1**	**0**	**0.5062**	**0.012**	**0.988**
164	531870	4171885	Southern	0	1	0	0.506	0.035	0.965
165	531815	4171710	Southern	0	1	0	0.5064	0.027	0.973
166	531814	4171710	Southern	0	1	0	0.506	0.004	0.996
167	531748	4171738	Southern	0	1	0	0.5063	0.014	0.986
168	**531743**	**4171730**	**Southern**	**0**	**1**	**0**	**0.5065**	**0.012**	**0.988**
169	531742	4171730	Southern	0	1	0	0.5061	0.011	0.989
170	531685	4171938	Southern	0	1	0	0.5066	0.025	0.975
171	531684	4171938	Southern	0	1	0	0.5061	0.226	0.774
172	531566	4172110	Southern	0	1	0	0.507	0.009	0.991
173	**531566**	**4172154**	**Southern**	**0**	**1**	**0**	**0.5074**	**0.043**	**0.957**
174	531562	4172133	Southern	0	1	0	0.5073	0.059	0.941
175	**531559**	**4172185**	**Southern**	**0**	**1**	**0**	**0.5066**	**0.099**	**0.901**
176	**531493**	**4173185**	**Southern**	**0**	**1**	**0**	**0.5067**	**0.012**	**0.988**
177	531419	4209075	Southern	0	1	0	0.5065	0.005	0.995
178	**531418**	**4209075**	**Southern**	**0**	**1**	**0**	**0.5065**	**0.007**	**0.993**
179	**530458**	**4180247**	**Southern**	**0**	**1**	**0**	**0.506**	**0.017**	**0.983**
180	530457	4180247	Southern	0	1	0	0.5064	0.008	0.992
181	530437	4172576	Southern	0	1	0	0.5061	0.224	0.776
182	530436	4172576	Southern	0	1	0	0.5065	0.052	0.948
183	**530299**	**4220164**	**Southern**	**0**	**1**	**0**	**0.5065**	**0.018**	**0.982**
184	**529929**	**4166452**	**Southern**	**0**	**1**	**0**	**0.5071**	**0.035**	**0.965**
185	529928	4166452	Southern	0	1	0	0.506	0.107	0.893
186	529909	4218670	Southern	0	1	0	0.5068	0.012	0.988
187	529908	4218670	Southern	0	1	0	0.5067	0.028	0.972
188	529849	4173024	Southern	0	1	0	0.5069	0.005	0.995
189	**529415**	**4217077**	**Southern**	**0**	**1**	**0**	**0.507**	**0.016**	**0.984**
190	**529027**	**4174820**	**Southern**	**1**	**0**	**2.954**	**0.0004**	**0.151**	**0.849**
191	528909	4174902	Southern	0	1	0	0.5064	0.166	0.834
192	528738	4214679	Southern	0	1	0	0.5071	0.069	0.931
193	528737	4214679	Southern	0	1	0	0.5067	0.027	0.973
194	**527664**	**4215871**	**Southern**	**0**	**1**	**0**	**0.5061**	**0.013**	**0.987**
195	**527635**	**4197230**	**Southern**	**0**	**1**	**0**	**0.5063**	**0.009**	**0.991**
196	527551	4209951	Southern	1	0	1.667	0.0021	0.149	0.851
197	526791	4182664	Southern	0	1	0	0.5067	0.054	0.946
198	**525886**	**4202603**	**Southern**	**0**	**1**	**0**	**0.5064**	**0.024**	**0.976**
199	**525470**	**4201294**	**Southern**	**0**	**1**	**0**	**0.5063**	**0.022**	**0.978**
200	**525467**	**4190073**	**Southern**	**0**	**1**	**0**	**0.5063**	**0.059**	**0.941**
201	**524836**	**4181942**	**Southern**	**0**	**1**	**0**	**0.5083**	**0.062**	**0.938**
202	**524413**	**4200169**	**Southern**	**0**	**1**	**0**	**0.5065**	**0.05**	**0.95**
203	**522317**	**4185104**	**Southern**	**0**	**1**	**0**	**0.5067**	**0.007**	**0.993**